PURDUE UNIVERSITY
OCT 16 1957
LIBRARY

820.5
W 885

ain Floor

VOL. 13 NO. 1 APRIL 1957

WORD

JOURNAL OF THE LINGUISTIC CIRCLE OF NEW YORK

EDITORS

WILLIAM G. DIVER ANDRÉ MARTINET URIEL WEINREICH

Assistant Editor: STANLEY LAMPACH

CONTENTS

On the Description of Phonic Interference ..... *Uriel Weinreich* 1
On the Relative Tractability of Morphological Data ..... *Paul L. Garvin* 12
Acoustic Cues for the Perception of Initial /w j r l/ in English ..... *J. D. O'Connor, L. J. Gerstman, A. M. Liberman, P. C. Delattre, F. S. Cooper* 24
On Discreteness and Continuity in Structural Dialectology ..... *Edward Stankiewicz* 44
Note on the Development of the Modern French Vowel System ..... *Alexander Hull* 60
The Phonology of the Word in Modern Standard Mongolian ..... *D. G. Stuart & M. M. Hallod* 65
The Phonetic Treatment of the Arabic Loanwords in Ethiopic ..... *Wolf Leslau* 100
Classifiers in Vietnamese ..... *Nguyen Đin Hoa* 124
Deux systèmes de vocabulaire parallèle à Madagascar ..... *Joseph Verguin* 153

REVIEWS:

D. Catalan M. P., La escuela lingüística española ..... *Geneviève Corrard* 157
W. Fucks, Mathematische Analyse von Sprachelemente, Sprachstil und Sprache ..... *B. Mandelbrot* 159
M. Schlauch, The Gift of Language ..... *A. G. Haudricourt* 160
G. Devoto, Dizionari di ieri e di domani ..... *G. Bonfante* 161
F. Antkowski, La chronologie de la monophtongaison des diphtongues ..... *W. Diver* 162
H. Kronasser, Vergleichende Laut und Formenlehre des Hethitischen ..... *A. Martinet* 164
H. Berger, Zwei Probleme der mittelindischen Lautlehre ..... *J. Manessy-Guillon* 166
V. Moltoni, Gli influssi dell'osco sulle iscrizioni latine della Regio I ..... *W. Diver* 167
U. Westerbergh, Chronicon Salernitanum ..... *R. L. Politzer* 170
M. Regula, Historische Grammatik des Französischen, Formenlehre ..... *A. Martinet* 172
H. Gneuss, Lehnbildungen und Lehnbedeutungen im Altenglischen ..... *E. V. K. Dobbie* 172
D. Jones, An Outline of English Phonetics & Phonetic Readings in English ..... *M. D. Berger* 175
D. Jones, Everyman's English Pronouncing Dictionary ..... *A. Martinet* 177
M. Niedermann, Balto-Slavica ..... *W. Schmalstieg* 178
G. Herne, Die slavischen Farbenbenennungen ..... *G. Y. Shevelov* 179
S. Bąk, Gwary ludowe na Dolnym Śląsku ..... *G. Y. Shevelov* 189
M. Rudnicki, Językoznawstwo polskie w dobie oświecenia ..... *G. Y. Shevelov* 192
Z. Stieber, Rozwój fonologiczny języka polskiego ..... *Z. Folejewski* 193
Z. Sobierajski, Gwary kujawskie ..... *Z. Folejewski* 195
L. Moszyński, Geografia niektórych zapożyczeń niemieckich w staropolszczyźnie. *Z. Folejewski* 196
R. Olesch, Zur Mundart von Chwalim in den früheren Grenzmark Posen-Westpreussen ..... *A. Martinet* 196
Ketica, Materialien aus dem Kethschen oder Jenisseiostjakischen ..... *R. Austerlitz* 197
L. Kettunen, Über die Wortfolge im Finnischen und Urfinnischen ..... *R. Austerlitz* 199
S. Egerod, The Lungtu Dialect ..... *A. G. Haudricourt* 200
A. Capell, A New Approach to Australian Linguistics & L. J. Luzbetak, Middle Waghi Phonology ..... *A. G. Haudricourt* 201

Published by
THE LINGUISTIC CIRCLE OF NEW YORK
New York, N. Y.

LINGUISTIC CIRCLE OF NEW YORK

Cercle Linguistique de New York

PRESIDENT : **Robert A. Fowkes**, New York University and Columbia University

VICE-PRESIDENTS : **Elliott Van Kirk Dobbie**, Columbia University
André Martinet, Sorbonne and Ecole libre des Hautes Etudes
Joshua Whatmough, Harvard University

SECRETARY and TREASURER : **James Macris**, Rutgers University

EXECUTIVE COMMITTEE : The preceding and

Marshall D. Berger, College of the City of New York
William G. Diver, Columbia University
Eugene Dorfman, University of Washington
Paul L. Garvin, Georgetown University
Roman Jakobson, Harvard University
Judah A. Joffe, Ecole libre des Hautes Etudes
Alphonse G. Juilland, University of Pennsylvania
Fredrick Jungemann, Queens College and Columbia University
Mel Most, New York
Lawton P. G. Peckham, Columbia University
Jeanne V. Pleasants, Columbia University
Morris Swadesh, Universidad Nacional Autónoma de México
Uriel Weinreich, Columbia University
Francis J. Whitfield, University of California

HONORARY PRESIDENT : **Henri F. Muller**, Columbia University and Ecole libre des Hautes Etudes

ADVISORY COMMITTEE FOR FOREIGN COUNTRIES :

Alfredo Barrera Vázquez, Mexico
Émile Benveniste, France
Giacomo Devoto, Italy
J. R. Firth, England
Henri Frei, Switzerland
J. Duchesne-Guillemin, Belgium
B. Havránek, Czechoslovakia
Louis Hjelmslev, Denmark
Bertil Malmberg, Sweden
J. Mattoso Camara, Jr., Brazil
†**Paul Menzerath**, Germany
Anton Reichling, The Netherlands
James St. Clair-Sobell, Canada
Alf Sommerfelt, Norway
Antonio Tovar, Spain

WORD, the journal of the Linguistic Circle of New York, is published three times a year.

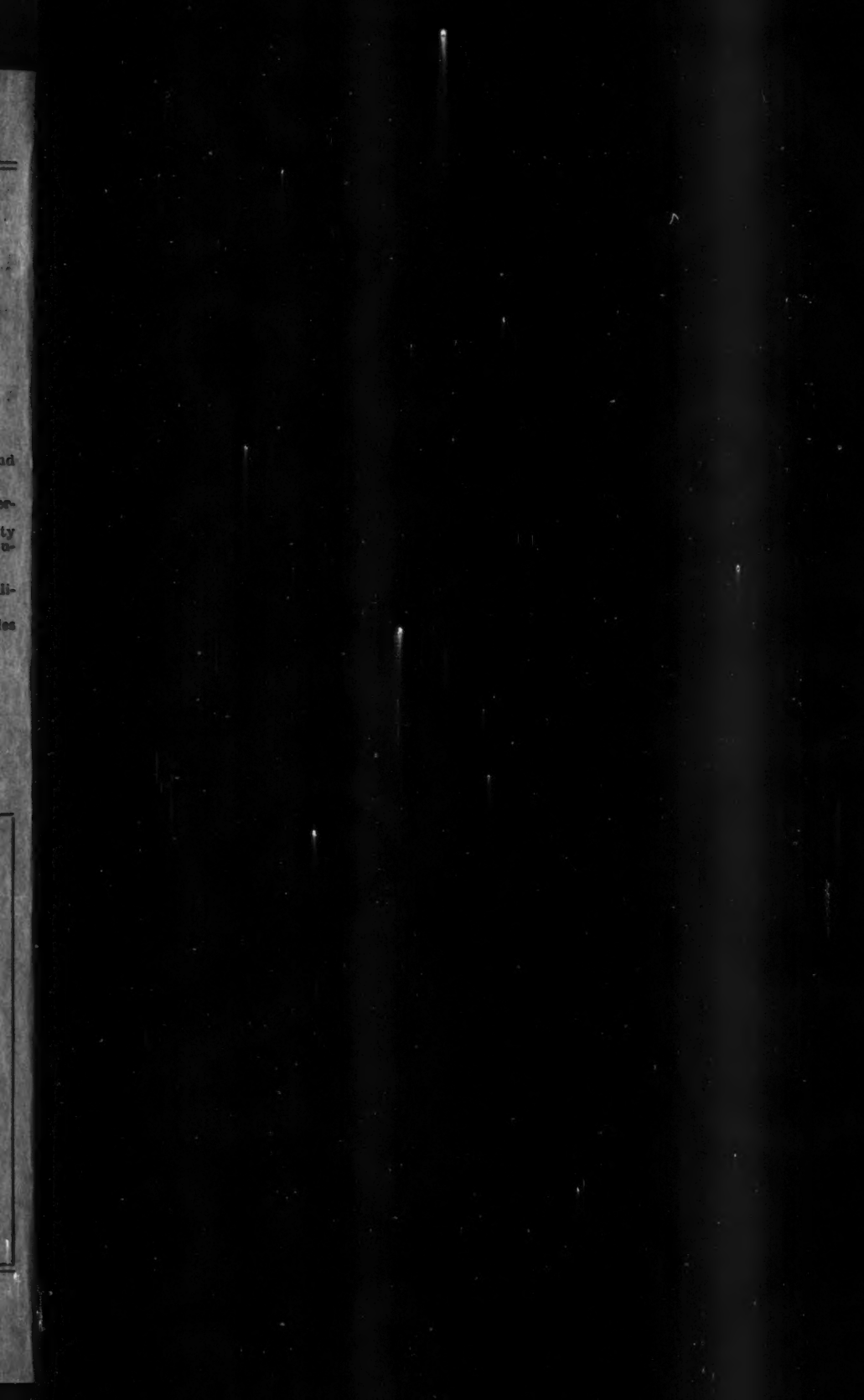
Subscription to WORD is $5.00 per year, this including membership. Dues for membership in the Circle without subscription to the journal are $1.50 a year. Single numbers of the journal to 1951 (Volume 7) inclusive, $1.50 ; single numbers from 1952 (Volume 8), $2.00 ; complete collection from Volume 1 through Volume 12, $53.00 (this including *Slavic Word*, Nos. 1 to 4, but not the monographs).

Applications for membership, subscriptions to WORD, and orders for back numbers should be sent to the treasurer: **James Macris**, Department of English, Rutgers University, New Brunswick, N. J.

Exchange journals, books for review or listing, and **manuscripts of reviews** should be sent to **André Martinet**, 3, place de la Gare, Sceaux (Seine), France. **Manuscripts of articles and monographs** should be sent either to **André Martinet** or to **Uriel Weinreich** or **William G. Diver**, Philosophy Hall, Columbia University, New York 27, N. Y.

The editors of the journal are jointly responsible for its contents. Authors of articles are allowed 100 reprints gratis and may get additional copies at cost. **Up to twenty free reprints of reviews are sent upon request.** Apply to the editor in charge of reviews.

The agent of WORD for Europe is Librairie Klincksieck, 11, rue de Lille, Paris, 7e.



WORD

VOLUME 13 *April, 1957* NUMBER 1

ON THE DESCRIPTION OF PHONIC INTERFERENCE

URIEL WEINREICH

In Chapter 2.2 of my book, *Languages in Contact* (New York, 1953; hereinafter *LiC*), I attempted to outline a plan for the description of phonic interference in its many aspects. That attempt was, I believe, not completely successful. The revised plan proposed here is based on further consideration of the problem, in which I have benefited particularly from discussions with Messrs. William J. Nemser (Columbia University) and Einar Haugen (University of Wisconsin) and from the latter's recently published work, including his able and constructive review of my book.[1]

By phenomena of INTERFERENCE I mean those instances of deviation from the norms of a language which occur in the speech of a bilingual as the result of his familiarity with another language, i.e. as a result of language contact (*LiC*, p. 1).

The problem of phonic interference concerns the manner in which a speaker perceives and reproduces the sounds of one language in terms of another. The language which causes the interference is called primary (*P*); the language which suffers the interference is designated secondary (*S*). *P*-Spanish means Spanish as the primary language in a contact situation; *S*-English, English as the secondary language. We are examining foreign "accent" in a broad sense—namely, language *S* spoken with a *P*-accent (*LiC*, p. 14).

[1] See Haugen's review of *LiC* in *Language* 30.380-8 (1954); "Problems of Bilingual Description," *Report of the Fifth Annual Round Table Meeting on Linguistics and Language Teaching* (Georgetown University Monograph Series on Languages and Linguistics 7), 1954, 9-19; "Problems of Bilingual Description," *General Linguistics* 1.1-9 (1955); *Bilingualism in the Americas: a Bibliography and Guide to Research*, Publications of the American Dialect Society nº. 26, 1957.

In reconsidering the presentation of *LiC*, it seems useful to make a forthright distinction between SYNTAGMATIC and PARADIGMATIC factors in phonic interference. The former pertain to relations between sounds in a sequence, i.e. in the spoken chain. The latter concern relations between sounds in the pattern, i. e. sounds which might occur at a given point in the spoken chain.

It is useful, furthermore, to distinguish between the bilingual's ANALYSIS of the material in the secondary language and his RENDITION of that material. One of the faults of the treatment in *LiC* is the inadequate separation of these two planes. Of course, it is possible to render unfamiliar sounds well by pure mimicry, yet such rendition without analysis hardly leads to productive speech. The bilingual's analysis in the present discussion is analogous to that process which enables us not merely to repeat a series of numbers like 1, 4, 9, 16, 25, ..., but to continue it consistently according to some hypothesis as to the formula.

Phonic interference may be due to factors on four levels:

I. Properly PHONIC FACTORS. This includes differences in the stocks of phonemes of the languages in contact, in the componential analysis and in the distributional patterns of their phonemes.

II. EXTRA-PHONIC FACTORS within the languages. Here we may consider cases of interference (or, on the contrary, suppression of interference) aimed at the avoidance of a particularly undesirable homophony, such as might cause serious ambiguity, unwanted allusions to taboo words in either language, and the like. These lexical considerations function as a control of interference in specific words and have nothing to do with the sound systems as such. They are analogous to the semantic factors which may occasionally upset an otherwise regular sound change. Possibly the influence of spelling on the pronunciation of the secondary language should also be classified in this category.[2]

III. EXTRA-LINGUISTIC FACTORS. These comprise the motivations for achieving intelligible, acceptable, or even native-like

[2] In *LiC* the influence of spelling is discussed only in connection with the form in which loan words from language *S* are integrated in language *P* (p. 28), but spelling as a model for the pronunciation of language *S* as such, is not dealt with. Yet in the pronunciation of *S*-English in India, for example, the spelling has been overwhelmingly influential. Cf. Ashok Kelkar's study of the English pronunciation of speakers of *P*-Marathi in a forthcoming issue of this journal.

speech that are present in a concrete speech situation or in the general sociocultural setting in which the language contact takes place. (*LiC*, § 1.3, pp. 25, 67, and chapters 3 and 4.)

IV. In any honest description we may also expect a residue of ERRATIC CASES of interference (see Haugen, *Language* 30.384). These are perhaps equivalent in status to the unaccountable kind of slips of the tongue that a unilingual speaker makes.

Here I will be concerned only with properly phonic factors (I).

Syntagmatic Factors

SEGMENTATION. A given sequence of speech sounds may be segmented differently by the speakers of different primary languages. Thus, a speaker of *P*-Hindi would segment the English sequences [ædhijr] *adhere* and [kləbhaws] *clubhouse* as containing the unit phonemes /d̲h, b̲h/. The [æ·] of English *bad* might be segmented as a cluster, /ææ/, from the point of view of *P*-Finnish. Similarly, the sequence [ts] at the end of English words might —morphological considerations apart—be analyzed as a unit phoneme by a speaker of *P*-German. True, if such phonemic metanalysis is possible without leaving a remainder of segments unaccounted for, the rendition of the *S* material by the foreigner may be quite native-like in this respect. But it is submitted here that the difference of analysis is a legitimate object of the linguist's interest even when rendition is not affected adversely.

FAMILIAR SEQUENCES. It is often observed that bilinguals, in mispronouncing words of *S*, adapt unfamiliar sequences of sounds to sequences more like those which occur in *P*. Thus, the sequence /ekst/ as in English *extra* is often simplified to /est/ by speakers of *P*-Italian, where the long cluster in not familiar (*LiC*, p. 23). The implication of this for the description of phonic interference is that the unit of interlingual identification should not invariably be the phoneme (*LiC*, pp. 7, 14), but occasionally a sequence of phonemes. On the other hand, in the interests of economical description no precise length of this sequence should be specified; it should simply be a possibility held open, like the possibility of occasionally equating not words, but stretches of words ("idiomatic constructions") in a bilingual dictionary. The whole matter of unfamiliar sequences is treated too casually in *LiC* (p. 23) as a structural factor which merely "favors" interference.

Obviously a sequence of sounds may be familiar in language *P*, but only in positions limited with respect to a morpheme or word boundary. Thus, *hn* is a sequence common in Ukrainian which does not occur anywhere in English; but the *ŋi-* of an Indonesian word is a sequence which, though it does not occur initially in English, is quite common word-medially *(singing)*. A sequence in an unfamiliar position, as we all know, often strikes us as no less exotic than a completely unfamiliar one. The relative value of total and "positional" familiarity in governing interference is capable of systematic investigation.

In general it is not yet clear, at the present stage of research, what methods may profitably be used in generalizing about classes of sequences. We may anticipate a technique by which it will be possible to say that *ðr-* in some language would be less liable to interference from *P*-English than *mr-* because, though neither cluster occurs initially in English, the former does approximate occurring English clusters more closely. After all, there may be holes in distributional patterns, too. Related to this point is the possibility of expressing certain aspects of the combinability of phonemes in quantitative terms, as established by the recent work of Herbert H. Paper.[3] His procedures will enable us, by comparing the distribution indices of pairs of equivalent phonemes of *P* and *S*, to make quantitative statements of difference of distributional structure of the two languages. However, the description of particular instances of interference will still have to refer to discrepancies in specific familiar sequences in the languages in contact.

It would seem that when we deal with unfamiliar sequences of *sub*phonemic segments (and perhaps in other cases, too), an analysis is possible in both syntagmatic and paradigmatic terms; cf. the chapter on the intersection of these two dimensions, p. 8 below.

A particular type of sequence discrepancy is that in which order alone is involved. Metatheses of sounds in the integration on unfamiliarly shaped loanwords are well documented in historical linguistics. It is also possible to conceive of cases in which a segment of sound in language *P* is distinctive but the place of its occurrence is not significant. This might lead to a mutilation of sound sequences of *S* such that all the phonemes are rendered,

[3] F. Harary and H. H. Paper, "Towards a General Calculus of Phonemic Distribution", to appear in *Language*, vol. 33, n°. 2 (1957).

but not in the right order. A real if somewhat marginal instance is that of a child which, on learning to pronounce *r* between a consonant and a vowel at the age of 22½ months, began, in polysyllabic words requiring *CrV*, to insert the new sound in any syllable at random; for a period, *professor* was likely to turn up half the time as *pofressor*, and *breadbox* alternated with *beadbrox*. Whether a corresponding phenomenon occurs systematically in non-child speech is open to question; at any rate many of us from our fieldwork know the feeling of identifying the constituent sounds of a word in a strange language, but not being sure of the order in which they were said.

Paradigmatic Factors

We consider first the paradigmatic aspects of the bilingual's ANALYSIS of a sound. It is convenient, as Haugen has shown, to deal with the allophone first, and the term 'sound' in the following paragraphs is to be interpreted as referring to an allophone, i.e. a contextual variant of a phoneme.

A sound of a given language *S*, e.g. the [ð] of Spanish, has many features: [ð] is 1. voiced (rather than voiceless), 2. oral (rather than nasal), 3. fricative (rather than occlusive), 4. apical (rather than medial or dorsal), 5. dental (rather than interdental or alveolar or postalveolar or cacuminal), 6. of a certain average duration (rather than longer or shorter than that), 7. non-pharyngalized (rather than pharyngalized), and so on. From the point of view of Spanish, some of these features, viz, 1, 2, and 4, are distinctive, while the rest are redundant, i.e. automatically predictable. In an analysis of the sound [ð] conforming to the phonemic pattern of a different language, say of *P*-English, feature 3 would also be distinctive. If *P* were Arabic, 6 and 7 might be distinctive in addition. Whenever a bilingual classifies a feature of sound as distinctive which in the original *S*-language is redundant, we may speak of OVER-DIFFERENTIATION. In the opposite case—as when a speaker of *P*-Spanish analyzes the fricativity of an intervocalic English [ð] as redundant—we have UNDER-DIFFERENTIATION.

The advance represented by this statement compared to that in *LiC* (p. 18) lies in the realizations that (a) over- and under-differentiation are matters of distinctive features rather than of phonemes, since conceivably in the bilingual's analysis of a single phoneme there may be both an under-differentiated feature and

an over-differentiated one: a speaker of *P*-French might consider the lip rounding of [ó] in Russian [tʹótʹa] 'aunt' as distinctive (over-differentiation) and its stress as redundant (under-differentiation); (b) over- and under-differentiation are matters of analysis rather than rendition; even under-differentiation may lead to impeccable renditions. In the previously cited example, the speaker of *P*-Spanish may under-differentiate the feature of fricativity which distinguishes English /d/ from /ð/, but intervocalically he would render a satisfactory [ð] nevertheless.

A situation of a special sort arises when language *P* classifies a distinctive feature of a given sound of *S* as redundant, but it does possess this distinctive feature in others of its own sounds. Thus the voicing and the laxness of *S*-Romansh /ž/ are redundant from the point of view of *P*-Schwyzertütsch phonemics, but the laxness at least is distinctive in such other Schwyzertütsch sounds as /b, d, v/, etc. This may be conducive to a more correct analysis of Romansh /ž/ by bilinguals than, say, of Romansh /ś/, whose feature of palatality is not distinctive for *any* sound of *P*-Schwyzertütsch. It is this which was mentioned as a factor of paradigmatic structure that inhibits interference, and was related to the concept of holes in the pattern, in *LiC* (p. 23). We may now state more precisely that under-differentiation of a distinctive feature may affect all, or only some, of the phonemes of the *S*-language which are characterized by that feature.

If in a language *S* a certain distinctive feature is associated with great regularity with a certain redundant feature (e.g. backness and lip-rounding of [u]), and if another language, *P*, manifests the same association except that the classification of the two features as distinctive and rendundant is reversed, the two sounds may be effectively identified astraddle of the languages, producing an elementary case of reinterpretation of features (cf. also p. 9 below). A workable reinterpretation should lead to renditions which are faultless with respect to the features concerned.

When we pass on from questions of analysis to those of RENDITION, our point of view shifts. What we care about here is whether a feature (be it distinctive or not) is always present in a sound, or whether it occurs only occasionally as a matter of free variation, or never. If the presence of a feature in a sound of *S* is a matter of free variation, any treatment of it by the rules of another language, *P*, ought to yield faultless renditions, whatever those rules may be, since by definition the presence or absence of the

feature does not matter for *S*. Also, if a feature is always present in the equivalent sound of *P*, the rendition will be satisfactory, whatever the regularity or irregularity of this feature in the corresponding sound in *S*. On the other hand, errors of varying predictability may be expected in other cases. For a given feature of the equivalent sounds of languages *P* and *S*, the results are as follows if the feature is present:

IN THE SOUND OF LANGUAGE *S*	IN THE SOUND OF LANGUAGE *P*		
	Always	*Sometimes (free variation)*	*Never*
Always	+	Unpr. err.	Pred. err.
Sometimes (free variation)	+	+	+
Never	Pred. err.	Unpr. err.	+

+ = Satisfactory rendition
Pred. err. = Predictable errors
Unpr. err. = Unpredictable errors

For example, in *S*-English, final /p/ is sometimes exploded, at other times not; the feature of release is a matter of free variation. In the English speech of a bilingual in whose *P*-language the release of final /p/ is regular, a released [p] would perhaps be more frequent than in the speech of an English unilingual, but there would be no perceptible interference. In the reverse case, in which English is the *P*-language, there would be unpredictable errors on this point, since the bilingual would treat the release of [p] as a matter of free variation contrary to the rules of the *S*-language. Examples of the other possibilities are easily invented.

At this point it is quite feasible to proceed from allophones (contextual variants) to a statement of how the PHONEME AS A WHOLE is rendered by the bilingual. We then necessarily obtain an additional class of errors, namely those which are predictable or unpredictable of the phoneme in a particular context. For example, if a speaker of a language *P* in which voicing of /b/ is a matter of free variation, were to render forms in *S*-German (where voicing of /b/ is always present in intervocalic position), we might expect errors in the rendition of /b/ in some contexts (namely intervocalically) but not finally.

Haugen's approach to the description of interference differs from the one proposed here chiefly in that, while he occasionally comments on the relative importance of a feature to a given language (*Norwegian Language in America*, p. 432), he does not take the componential analysis of all phonemes as the basis of description. But whatever role is assigned to the distinctive feature in analyzing the allophones of a given language, when the cross-language equivalences between allophones have been established—by the present method or by any other—there can be no objection to the adoption of Haugen's procedure for classifying them into "diaphones" (interlingually identified phonemes) of various types: convergent and divergent, simple, compound and complex. (See Haugen's latest formulation in *Bilingualism in the Americas*, § 3.3.3.)

Intersection of Syntagmatic and Paradigmatic Factors

There are cases of interference in which syntagmatic and paradigmatic factors intersect.

One type of intersection occurs when a phoneme of *S* has an allophone which is subject to further segmentation in terms of language *P*. Thus, Tamil /e/ has, in word-initial position, the allophone [je]. Faced with the English sequence [jel] *yell*, the speaker of *P*-Tamil would resegment it so that the first two segments would together constitute the word-initial variant of his /e/. Faced with /el/ *ell*, he would probably note a strange initial allophone of /e/ and, in rendering it, he would supply the additional "suballophonic" segment [j]-. From a syntagmatic point of view, we have a difference in segmentation in the two languages. Paradigmatically, we might speak of a discrepancy in the regularity with which the "feature" of *j*-onset occurs—never in English, always initially in Tamil. A bilingual's treatment of discrepant functioning of subphonemic glides between vowels, epenthetic vowels and consonants, and similar phenomena would likewise involve a choice of syntagmatic and paradigmatic terms in the description of the interference.

Another category of intersection is that in which a special segment is inserted into *S* forms by the bilingual to manifest a feature which in *P* does not occur distinctively in a particular sound. Thus, a type of *P*-Ukrainian in which a distinctively voiceless /f/ does not occur causes bilinguals to render [xv] for the *f*-sound of secondary languages. The bilingual unwraps, as it

were, the bundle of simultaneous features which constitute [f] and sets up a special segment for the display of voicelessness. Conversely, the rendition of a labialized [k^w] by a speaker of a *P*-language which has distinctive labialization, in lieu of the sequence [kw] of an *S*-language which does not, trades a segment for a simultaneous distinction.

Perhaps the most interesting variety of intersection occurs when a distinctive feature is transferred to another, existing segment of *S*: it is taken out of one bundle of features and wrapped in another. Thus the Romansh sequence [ɛs·] as in *Messa* 'Mass' contains a consonant which is distinctively long (or geminated) and a vowel which is redundantly short, but whose shortness is required by what follows. The same sequence may be interpreted in terms of a neighboring type of Schwyzertütsch as consisting of a vowel which is distinctively short and a consonant redundantly, but automatically, long. This is what was called "reinterpretation of features" in *LiC* (pp. 18f.), but it is a more complex case than the elementary one outlined above (p. 6), since a phoneme boundary is crossed here. The features of each of a pair of phonemes are reinterpreted as to distinctiveness, and in addition, the two allophones continue to imply each other despite the reinterpretation. Another example is the opposition, in *S*-Arabic, of /aq—ak/ [ɑq—ak]; a speaker of *P*-French might consider the frontness of the vowel distinctive rather than the pharyngalization of the consonant: /ɑk—ak/. The parallelism with the concept of the preservation of useful sound features will be apparent to all who are familiar with diachronic phonemics.

Inference in Prosodic Features

Interference in prosodic features was treated in *LiC* in a somewhat too fragmentary manner (pp. 17, 18, 38f.). In a more systematic analysis, several types of stress system and the possible interferences should be examined. We leave aside the bilingual's problems that bear on the use of stresses and stress contours as meaningful, morpheme-like units; only the distinctive function of stress is touched on here.[4]

We may consider languages which have culminative stress,

[4] On the conflicting functions of stress in a single language, see my paper, "Stress and Word Structure in Yiddish," in *The Field of Yiddish: Studies in Language, Folklore and Literature*, New York, 1954, 1-27.

i.e. a stress which occurs in a determined but unpredictable place in each word; demarcative stress, i.e. a stress which occurs automatically in a determined place in each word in a position related to the beginning or the end of the word (final, pre-final, initial, etc.); and non-distinctive stress, i.e. a stress which plays no role in identifying or segmenting words and which may occur at any point of a word, without any necessary consistency. Examples of languages with culminative stress are English, Russian; demarcative stress occurs in Latin, Czech, etc.; non-distinctive stress is the rule in French.

If the *P*-language has culminative stress, we may expect the bilingual to pay attention to the stresses of the *S*-language and to render them correctly. If he has not attained the automaticity of stress placement which is possible in an *S*-language with demarcative stress, or if he places stresses consistently in an *S*-language in which they are not distinctive and need not be consistently placed, there is a lack of economy analogous to the over-differentiation discussed under the heading of paradigmatic factors; but the rendition may be perceived as faultless.

If the bilingual is not accustomed, by the rules of *P*, to place the stress on a separately determined part of each word, errors in rendition may be expected. When the required place of the stress of a form in *S* happens to coincide with the automatic place of demarcative stress in *P*, the rendition may be adequate by accident; otherwise there will be faulty stressing. For example if a speaker of *P*-Czech (demarcative initial stress) were to apply his habits to *S*-English, he would happen to get *discipline* right but *disciple* wrong (**disciple*). A speaker of *P*-French, if he treated the stress of *S*-English as if it were non-distinctive, would sometimes happen to render words correctly, but according to no predictable formula. The same word might be produced with initial stress once and final stress at another time. Thus, it is clear that there will be considerable errors in the rendition of *S* if the stress system of *S* is more rigorous than that of *P*(the progression non-distinctive—demarcative—culminative being treated as increasing in rigor). Also, the place of likely errors may be predictable if either *P* or *S* or both operate with demarcative stress.

This is a mere outline of a plan for the description of interference in stress. The bilingual's treatment of secondary and lesser stresses should be analyzed in a similar way. The demarcative

[8] See André Martinet, "Accents et tons," *Miscellanea phonetica* 2.13-24 (1954).

type of stress may be divided further into word-demarcative and stem-demarcative; and still other refinements may be introduced.[5] But no purpose would be served by simply listing all the possible categories of interference; the analogy with the description of segmental-phoneme interference makes it clear how this classification could be extended.

Interference in other prosodic features, particularly tone systems of various types, remains to be described on related principles.

Finally, the analogy between interference in segmental phonemes and in prosodic features could be pursued further by introducing such concepts as "allophones of stress." But no terminological exercise is likely to make the already apparent parallelisms any clearer, nor to eliminate the ultimate dissimilarities between segmental phonemics and prosody.

Columbia University.

ON THE RELATIVE TRACTABILITY OF MORPHOLOGICAL DATA*

PAUL L. GARVIN

0. It has been customary among linguists for some time now to tell students that "there is no such thing as a difficult language," and that the difficulty of learning a language is a subjective matter and a question of the learner's native language. At the same time, in discussing with each other the problems of their field work, linguists would readily evaluate the relative tractability of their material; on this level of the argument, linguists would agree that, for instance, some North American Indian languages present considerable difficulty to morphological analysis, while on the other hand the languages of, say, the Malayo-Polynesian family are relatively easy to analyze morphologically. These impressionistic judgments show a consistency which strongly suggests that the matter lends itself to systematic treatment, leading perhaps to a set of criteria for validating such judgments with regard to morphological tractability.

The procedures of morphological analysis can, roughly speaking, be summed up in the following four major steps:

(1) isolating minimum forms (morphs);

(2) assigning morphs to morphemes as their allomorphs (morpheme alternants or morphophonemic variants);

(3) ascertaining morpheme distributions and establishing morpheme classes;

(4) exhausitve listing of morphemes and morpheme classes.

The ease or difficulty with which each of these steps can be completed depends on the readiness with which the required operations can be applied to the material of a given language; languages vary in this respect.

* A first version of this paper was read at the meeting of the Linguistic Society of America in Philadelphia, December 28, 1956.

1. ISOLATING MORPHS. For operational purposes, a minimum meaningful form can be redefined as the smallest separable stretch of form with a consistent identifiable piece of meaning. The first step in morphological analysis thus involves two sets of operations: separating forms, and identifying the presence of a consistent piece of meaning for each.

1.1. SEPARATING FORMS. This is essentially accomplished by two closely related procedures: substitution and dropping. By dropping I mean the elicitation or comparison of examples such that a given partial present in one example will, all other things being alike, be absent in another example.

(a) The substitution procedure is applied most readily when paradigmatic sets can be easily elicited, as in the following examples from Kutenai:[1]

Set I *huc ʔinaxe·* 'I'm going there'
hinc ʔinaxe· 'you (sg.) are going there'
c ʔinaxe· 'he, etc., is going there'
huc ʔinaxałaʔne· 'we are going there'
hinc ʔinakiłne· 'you (pl.) are going there'
c ʔinaxe· 'they are going there'

Set I unequivocally yields as separable forms *hu-*, *hin-*, and *ø-* (zero), for first, second, and third person respectively. It also suggests forms *-ałaʔ-* and *-kił-* as plural markers for first and second person respectively.

Set II *hucxałcʔinaxe·* 'I shall go there'
huqa·kiłcʔinaxe· 'I used to go there'
husiłcʔinaxe· 'I am on my way there'
hucxałsiłcʔinaxe· 'I'll be on my way there'

Set II yields the separable forms *-cxał-* and *-qa·kił-* for future and past respectively, and form *-sił-* with something like continuative meaning.

(b) The dropping procedure is applied most readily when the elicitation of paradigmatic sets reveals the presence of clearly bound and clearly free forms, as in the following examples from Ponapean:[2]

[1] Spoken by the Kutenai Indians of northern Idaho, northern Montana, and southeastern British Columbia.

[2] Spoken by the natives of Ponape, Eastern Caroline Islands, Micronesia.

Set I *ikilaŋ* 'I see'
kəkilaŋ 'you (sg., non-honorific) see'
ɛkilaŋ 'he, etc., sees'
kilaŋ 'see'

Set I yields the bound forms *i-*, *kə*, and *ɛ-*, for first, second, and third person, respectively, and the free form *kilaŋ*.

Set II *áramaso* 'that man'
áramasaka 'people'
áramas 'man'

Set II yields the bound forms *-o*, *-aka* for demonstrative and plural respectively, and the free form *áramas*.

(c) The application of both procedures to the sets presented above is obviously "easy." The reason for this "ease" is equally obvious: form boundaries are clearcut, and the separation points are unmistakable.

The difficulties of separating forms then must arise when boundaries are not clearcut. Let me exemplify this with Ponapean examples:

ò·lɛnpó·npey 'man of Ponape'
mɛó·l it's a man'
mè·npó·npey 'Ponapeans'
olɛnwáy 'foreigner'
mè·nwáy 'foreigners'

In the above substitution set, the forms *ò·l-* or *ol-* 'man', *-ɛn-* 'of', *-pó·npey* 'Ponape', *-wáy* 'abroad' (assuming that *olɛnwáy* by analogy with *ò·lɛnpó·npey* stands for something like 'man of abroad'), can be clearly separated out, and the form *mè·n-* initially appears indivisible, tempting the analyst to include it in his list as a single morph. It is only after considerable additional analysis that it becomes clear that the form *mè·n-* is the result of the contraction *mɛ+ɛn;* the form *mè·n-* as it stands in the elicited set does not immediately suggest a division in two. Even after the division has been accomplished, the boundary between the two components cannot be clearly drawn, since only the initial and final consonants of the form *mè·n-* can unequivocally be assigned to the underlying components *mɛ* and *ɛn*, the medial vowel *e·* is the indivisible result of the contraction. Such phonemic units shared by two morphs, and without allowing the clearcut placement of a boundary within, I have called *ambimorphemic*,

and the partial phonemic fusion of two morphs I call *morpheme overlap*.

(d) An extreme instance of morpheme overlap exists in the case of portemanteau forms, where a single separable (and indivisible) form fills a place in a paradigm otherwise occupied by two separable forms which may or may not be overlapping, as in the French sequences of preposition, article, noun:

ovwazę
dyvwazę
parləvwazę
avɛkləvwazę

(e) The separation difficulty can thus be related to the existence of morpheme overlap, and allows the formulation of the following rule: *the difficulty of separating morphs increases with the amount of morpheme overlap present in the language.*

1.2. *Identifying meaning.* Operationally, linguistic meaning can be handled by translation or, if the analyst is a native or near-native speaker of the language, by paraphrase. The presence of a consistent piece of meaning in a separated stretch of form can thus be ascertained by checking the translations (or paraphrases) relevant to it.

(a) The simplest case of identifying the presence of meaning exists when it is possible to elicit clearcut, consistent individual translations of each separated form, as is the case for Set I under (b) in **1.1.**

Informants will readily translate each separated form, and do so consistently: *i-* as 'I', *kə-* as 'you (sg., non-honorific)', *ɛ-* as 'he, etc.', *-kilaŋ* as 'see'. Note that such forms may include bound forms as well as free forms.

(b) Sometimes even good bilingual informants will readily identify a form as separate and having meaning, but will be hesitant and inept in giving a translation. This happens particularly frequently when the forms in question do not correspond to readily available forms in the other language.

Thus, Kutenai informants will follow the linguist in separating out forms such as *ʔał* 'indeed', *pał* 'really', *miqsa·n* 'however', but will have some difficulty in giving a translation. They will, however, agree that these forms do have separate meanings.

(c) At times, informants will identify forms as separate and readily volunteer definite individual translations, but the translations will, upon closer investigation, lack consistency and appear contradictory to the linguist. Thus, Visayan[3] informants will usually translate as follows:

gibúhatku ʔaŋbaláy 'I'm building the house'
nakiʔtaʔku ʔaŋbaláy 'I'm seeing the house'

They will furthermore volunteer to translate the morph *-ku* as 'I', and the morph *ʔaŋ-* as 'the'. Later extensive analysis of the verb pattern reveals, however, that these translations are deceptive: the examples involve passive verb forms, and should "properly" be translated literally as 'the house was built by me', 'the house could be seen by me', with the morph *-ku* translated as 'by me' and the morph *ʔaŋ-* constituting a subject marker.

(d) Very frequent are the cases where translations can be obtained only for the several complete forms of a paradigm, and the translations of the separable morphs have to be identified by matching minimal differences in form against minimal differences in translation. Thus, in the Kutenai paradigm cited as Set I under (a) in **1.1**, individual translations cannot be obtained for the easily separable forms *hu-* for first person, *hin-* for second person, etc., but the identification of their meanings follows readily from the translation of the examples as wholes.

(e) Finally, I would like to cite cases in which the separation of forms appears quite obvious, but where the assignment of meanings presents considerable difficulties. Here belong the famous English sets such as :

receive	*revert*	*retain*
conceive	*convert*	*contain*
deceive		*detain*
perceive	*pervert*	*pertain*
	subvert	

A casual inspection of the above collection will indeed reveal that each of these forms can by the substitution technique easily be divided into two components, yielding the suspected morphs *re-*, *con-*, *de-*, *per-*, *sub-*, *-ceive*, *-vert*, *-tain*. The difficulty arises

[3] A Philippine language.

when the presence of consistent meaning is to be identified: neither translation into another language, nor paraphrasing within English (nor, for that matter, recourse to the native speaker's intuition) yields a *consistent* piece of meaning for any of these separated forms. It is thus not by accident that these sets have been the subject of much controversy among linguists.[4]

(f) It is clear from the above that the identification of meaning involves a greater variety of problems (and hence of techniques) than the separation of forms, since meaning—as is well known—is less immediately apparent than form, and hence its presence must be ascertained by techniques that are increasingly inferential. This allows the formulation of the following rule : *the difficulty of identifying the presence of meaning in separated forms increases with the number of inferential steps required.*

2. Assigning morphs to morphemes. Two or more morphs can be assigned as allomorphs to the same morpheme if (1) their meanings fall within the same acceptably narrow range, (2) they are in free variation or complementary distribution. The two criteria are clearly interrelated: on the one hand, two morphs in the same distributional position will be considered in free variation only if their meanings can in no way be said to be in opposition; on the other hand, two morphs will be examined for complementary distribution only if they can be shown to have (or at least can be suspected of having) a significant feature of meaning in common.

The second step in morphological analysis thus involves two sets of operations which must be performed in close conjunction: collating morph meanings, and examining morph distributions. In testing for free variation, the sameness of distributional position is given and the non-opposition of meaning must be demonstrated. In testing for complementary distribution, the non-opposition of meaning is given (or assumed) and the exclusiveness of distributional conditions must be demonstrated.

2.1. Morphs in free variation. Phonemically unlike morphs can safely be considered in free variation if forms differing only by these morphs are consistently considered identical by native speakers.

(a) In Kutenai, such forms as *hu-cxaɬ-ʔik-ne·* or *hu-c-ʔik-ne·* 'I shall eat', *hin-cxaɬ-ʔin-ne·* or *hin-c-ʔin-ne·* 'you (sg.) will be',

[4] For a discussion in some detail, see Dwight L. Bolinger, "On Defining the Morpheme," *Word* 4.18-25 (1948).

are consistently said to "mean exactly the same"; they are often used interchangeably or given as repetitions of each other. The morphs *-cxal-* and *-c-* thus have to be considered in free variation; they are morphemic free variants in all distributions.

(b) In Visayan, a number of verb stems occur in either of two phonemic shapes when followed by a suffix; the verb forms in question are completely interchangeable:

gi-lisud-án or *gi-lisd-an* 'have difficulty with'
na-walá?-an or *na-wál?-an* 'lose'
gi-salipud-án or *gi-salipd-an* 'be hidden'

The morphs *-lisud-* and *-lisd-*, *-walá?-* and *-wál?-*, *-salipud-* and *-salipd-*, are morphemic free variants, but unlike the previous case, the free variation is restricted to one set of distributions only.

(c) In Czech, the two locative singular morphs *-ě* and *-u* appear to be in free variation after certain inanimate noun stems ending in a 'hard' consonant, such as:

bod-ě ∾ *bod-u* 'point (loc.)'
hlas-e ∾ *hlas-u* 'voice (loc.)'
jazyc-e ∾ *jazyk-u* 'tongue; language (loc.)'

A closer examination of the examples reveals, however that for instance, of the third pair *jazyc-e* is used more frequently in contexts where the stem is translated as 'language', and *jazyk-u* more frequently in some contexts, and exclusively in other contexts, where the stem is translated as 'tongue' in the literal sense. The free variation is thus not only distributionally restricted, but the coincidence of meaning of the morphs involved is not clearcut.

(d) The above cases can be summarized as follows: *The difficulty of assigning morphs to morphemes increases with the distributional restriction of the suspected free variation, and decreases with the clarity of the meaning equivalence.*

2.2. MORPHS IN COMPLEMENTARY DISTRIBUTION. Phonemically unlike morphs can safely be considered in complementary distribution if forms containing these morphs consistently share a feature of meaning assignable to the morphs in question, and are otherwise morphemically different.

(a) In Ponapean, such forms as *ɬółɔ·k* 'work' and *ɬɬó·ŋ-ki* 'work with' share the feature of meaning 'work', such forms as

sɛ́ylɔk 'travel' and *sɛylɔ́ŋ-ki* 'travel by' the feature of meaning 'travel', assignable to the morphs *łɔ́łɔ·k* and *łɔłɔ́·ŋ-*, *sɛ́ylɔk* and *sɛylɔ́·ŋ-* respectively; they are morphemically different, since the second of each pair contains the morph *-ki* 'by, with', the first of each pair does not. A great many such forms can be collected in Ponapean, showing that *k*-final morphs are replaced by *ŋ*-final whenever the following morph is *k*-initial. These sets of morphs are thus in complementary distribution under phonemic conditions. I shall call such allomorphs *morphophonemic variants*.

(b) In Kutenai, such forms as *hu-q'umne·-nała'-ne·* 'we are sleeping' and *hu-c'inax-ała'-ne·* 'we are going there' share the feature of meaning 'plural of first person' assignable to the morphs *-nała'-* and *-ała'-* (cf. (a) under **1.1.**); they are morphemically different, since one contains the morph *-q'umne·-* 'sleep' and the other the morph *-c'inax-* 'set out for'. A great many such forms can be collected in Kutenai, showing that when preceded by a certain set of morphs, the first person pluralizer morph will be *-nała'-*, when preceded by another set of morphs, it will be *-ała'-*. The two morphs are thus in complementary distribution under morphemic conditions; I call such allomorphs *morpheme alternants*.

(c) In Kutenai examples such as *hu-n-'upx-ne·-mał-ne·* 'we were talking to each other', *ø-cxał-cukał-e·-mał-ne·* 'they'll get married', *hin-anaq-niy-kił* 'if you sat down', *hu-n-'aqaciy-iy-ała'-ne·* 'we cut our hands', *ø-qa·c'ła'nxu-na-k-s-e·* 'a fallen tree was lying there', or *ci'kał-a-m* 'look at yourself', the morphs *-ne·-*, *-e·-*, *-niy-*, *-iy-*, *-na-*, *-a-* will, upon further analysis, be found to be in complementary distribution, and a common grammatical meaning of 'reflexivizer' can be assigned to them, even for cases where it is not obvious from the translations given above. It will then also be found that the choice between the morphs *-ne·-*, *-niy-*, *-na-* on the one hand, and between the morphs *-e·-*, *-iy-*, *-a-*, on the other, will depend on the same preceding sets of morphs as the choice between *-nała'-* and *-ała'-* in the examples cited above. Within the set of morphs *-ne·-*, *-niy-*, *-na-*, and within the set *-e·-*, *-iy-*, *-a-*, the choice is made in terms of one or several of the following morphs. In addition to this morpheme alternation, the morphs *-ne·-*, *-e·-* containing a long vowel are also subject to a morphophonemic variation affecting vowel length.[5]

[5] For details of this allomorphic variation, see my "Kutenai III: Morpheme Distributions (Prefix, Theme, Suffix)," *IJAL* 14.185-6 (1948).

(d) In English, the morphs *-s*, *-z*, *-iz*, *-in*, are generally considered to be in complementary distribution and to constitute some of the allomorphs of the plural suffix, as in *cats*, *dogs*, *horses*, and *oxen*. The choice between *-s*, *-z*, *-iz* on the one hand and *-in* on the other depends on the preceding morph, the choice within the set *-s*, *-z*, *-iz* depends on the preceding phoneme. We are thus dealing with two morpheme alternants, one of which in turn is made up of three morphophonemic variants. The situation appears to be clearcut so far. A difficulty arises, however, from the well-known fact that in such examples as *brothers* vs. *brethren* the two morphs *-z* and *-in* appear to be in oppositive distribution, since the morphemic environment is not obviously different (*bráðer-* and *bréðr-* may be considered allomorphs), and the difference in meaning between the two total forms is as easily assignable to the suffix morphs as to the stem morphs. As in the previously cited English examples, here again the set has given rise to a good deal of controversy.

(e) The above cases can be summarized as follows: *the difficulty of assigning morphs to morphemes increases with the complexity of distributional conditions and, as in the previous case, decreases with the clarity of the meaning equivalence.*

3. Ascertaining morpheme distributions and establishing morpheme classes. Once morphs have been assigned to morphemes, the distribution of morphemes as wholes, irrespective of allomorphic variation within each, is to be ascertained and they are to be grouped into classes on the basis of common distributional characteristics.

In the analysis of many languages (as well as in traditional grammar), such broad classifications as stems, prefixes, suffixes, have been intuitively accepted and have allowed a fairly good pragmatic assignment to classes. Even so, in instance after instance difficulties have arisen in terms of requiring a decision between assigning a morpheme to a class of, say, proclitics or one of prefixes, one of enclitics or one of suffixes. The well-known suggestions about the possible prefix status of French preverbal pronouns, or about the possible enclitic status of the Saxon genitive in English, serve to illustrate this. Since affixes are commonly designated as 'parts of the same word', and clitics as 'separate words', these difficulties have served to focus attention upon the problem of word boundaries, and upon the usefulness of the word as a distributional framework for morphemes.

Whereas in the case of the assignment of morphs to morphemes merely the sameness or not-sameness of the environment was a requisite criterion, in the case of the statement of morpheme distributions the environment has to be specified in detail in order to allow the precise statement of distributional differences between morphemes needed for morpheme classification. Such a precise statement of morpheme distribution requires as its proper distributional framework a larger morphemic unit, within the boundaries of which the occurrence of individual morphemes can be rigorously described.

The difficulty with which morpheme distribution can be stated will then increase in terms of the difficulty with which such a frame unit can be defined. If the definition of such a frame unit can be based on criteria of a different sort than the definition of any other larger morphemic unit, such a morphemic unit can properly be called a word. Conversely, if no such definition of a larger morphemic unit is possible, it can be said that the language in question has no words in our sense.

Examples of languages in which the word can be defined in a simple manner, and by criteria different from those for the definition of the phrase or clause, are Kutenai, Ponapean, and Turkish. In these languages, the word can be defined as a sequence of morphemes and morpheme clusters in fixed order. Morpheme clusters in turn are for these languages defined as sequences of morphemes which, within a longer chain of morphemes, are consistently replaceable by single morphemes.

In such languages, morphemes can be grouped into classes with relative ease in terms of the position they occupy with regard to word boundaries, or with regard to some specially defined morpheme or morpheme cluster within the word (such as a stem or theme), leading to a classification of morphemes into position classes.

I have tested this simple fixed-order definition of the word for French and English, with negative results, due to the absence of consistently ordered long chains of morphemes in these languages. I would assume that there are many additional languages in which the fixed-order criterion is inapplicable.

In languages such as French, Czech, and Polish, word boundaries have traditionally been identified with a phonological feature fixed within the contour, namely stress. This definition has been tested by me and found unviable, since it yields a large residue of multiword contours containing proclitics and enclitics not bounded

off by the required stress, and yet traditionally considered separate words. If, on the other hand, the phonological definition is carried out consistently and the traditional clitics are included into adjacent words as affixes, the statement of morpheme distributions within the word becomes unproductive, because it does not allow for the formulation of morpheme classes specific enough to lead to larger-unit classifications useful in syntax.

My attempts to define the word rigorously for French, English and Vietnamese by criteria other than phonology or fixed order, as well as by a combination of the two, have failed; I have come to the conclusion that a proper definition of the word, that is, a consistent one without unaccountable residue, will emerge only from a complete distributional reanalysis of the morphemic structure, using a properly defined temporary framework such as the utterance. The beginnings of such an analysis were made for English, yielding an interim classification of morphemes in terms of those occurrent in utterance-initial, medial and final, defining an utterance as a unit bounded by pauses, but barring hesitation forms and citation forms.

In summary, it can be said that *the difficulty of ascertaining morpheme distributions and of establishing morpheme classes increases with the difficulty of defining a proper distributional framework such as the word.*

4. EXHAUSTING THE NECESSARY LISTING OF MORPHEMES AND MORPHEME CLASSES. The criterion of exhaustiveness in the empirical principle[6] requires a complete listing of morphemes before a morphological analysis can be considered done.

For certain classes of morphemes, mainly those called inflectional or paradigmatic, exhaustive listing has never presented serious difficulties. For other classes of morphemes, primarily those called stems or bases, on the other hand, an exhaustive listing is ordinarily not even attempted, although partial lists may appear in lexical appendices. Let us for our purposes call the former classes of morphemes paradigmatic and the latter thematic, and examine this discrepancy in regard to listing.

It will be found that one viable definition of paradigmatic morphemes is that they belong to classes with limited membership, and of thematic morphemes that they belong to classes with

[6] See Louis Hjelmslev, *Prolegomena to a Theory of Language*, Francis Whitfield, transl. (Baltimore, 1953), p. 6.

unlimited membership. With this definition, the criterion of consistency—which in the empirical principle overrides the criterion of exhaustiveness—allows only a listing of the limited membership of paradigmatic classes, since an exhaustive listing of the unlimited membership of thematic classes is a contradiction in terms.

Thus, exhaustiveness will mean, for the paradigmatic classes complete listing of individual members of classes as well as of classes and subclasses, for the thematic classes complete listing of classes and subclasses, but mere exemplification—without complete listing—of individual members.

It will then be found that the number of paradigmatic classes and subclasses, as well as the size of the membership of each, varies from language to language, as does the number of thematic classes and subclasses.

Obviously, the more there is to list, the greater the difficulty of achieving exhaustiveness. Hence, *the difficulty of the exhaustive listing of morphemes and morpheme classes increases with the number and size-of-membership of paradigmatic classes and subclasses, and with the number of thematic classes and subclasses.*

5. The above criteria will allow at least a tentative evaluation of morphological raw material in terms of relative analytic tractability.[7] My experience in this respect strongly points to the fact that some of the well-known Indo-European languages are clearly less tractable morphologically than some of the "exotic" languages I have dealt with. This to me also explains some of the divergences and lack of agreement in regard to the morphological analysis of languages such as English and French.

The less tractable a language, the less will it be possible to apply to it the relatively cut-and-dried sequences of steps applicable to a more tractable language, and the more will it become necessary to devise new sequences of operations framed in terms of these intractable data.

Institute of Languages and Linguistics, Georgetown University.

[7] Joseph Greenberg, "A Quantitative Approach to the Morphological Typology of Languages," in *Method and Perspective in Anthropology*, Robert F. Spencer, ed (Minneapolis, 1954), pp. 192-220, has suggested some numerical criteria of use in typological classification. Some of these bear a direct relation to points covered in the present paper (thus, his index of synthesis to the separation difficulty treated in **1**); note also his very pertinent discussion of the definition of the word (216-8).

ACOUSTIC CUES FOR THE PERCEPTION OF INITIAL /w, j, r, l/ IN ENGLISH*

J. D. O'Connor,** L. J. Gerstman, A. M. Liberman,***
P. C. Delattre,**** and F. S. Cooper

In investigating the acoustic cues for the perception of speech it is usually convenient and often necessary to study only a relatively small group of phonemes in any single experiment. Thus, our own earlier research and that of our colleagues at Haskins Laboratories has dealt separately with the stop, nasal, and fricative consonants.[1] In the experiments to be reported here we have chosen to study /w, j, r, l/ as a group and to try to find the physical stimuli essential to the recognition of these phonemes in (absolute) initial position before vowels.

The selection of /w, j, r, l/ is necessarily arbitrary, at least in some degree; it is not random, however, since these phonemes have certain articulatory and distributional properties which tend to set them off from the other English consonants. From an articulatory standpoint, they differ from the constrictive consonants (stops, fricatives) in the degree of oral stricture present,[2] and

* This research was supported in part by the Carnegie Corporation of New York and in part by the Department of Defense in connection with Contract DA49-170-sc-1642.

** Also, University College, London.

*** Also, University of Connecticut.

**** Also, University of Colorado.

[1] A. M. Liberman, P. C. Delattre, F. S. Cooper, The role of selected stimulus-variables in the perception of the unvoiced stop consonants, *American Journal of Psychology*, 65.497-516 (1952); A. M. Liberman, P. C. Delattre, F. S. Cooper, L. J. Gerstman, The role of consonant-vowel transitions in the perception of the stop and nasal consonants, *Psychological Monographs*, 68. No. 8, Whole No. 379 (1954); Katherine S. Harris, Cues for the identification of the fricatives of American English, *Journal of the Acoustical Society of America*, 26.952 (1954).

[2] The majority of the allophones of /w,j,r,l/ are voiced oral resonants, and the occurrence of voiceless fricative allophones of all four phonemes in the same phonetic environment, i.e., after a voiceless stop or fricative in the same syllable, whilst no doubt connected with their phonemic distribution, is an added mark of their coherence as a class.

from the nasals by their oral articulation. On the basis of distributional characteristics, too, /w, j, r, l/ can be distinguished as a group from other sub-classes of English consonants. Their most obvious distinguishing mark is that they, and they alone, can constitute the third member of an initial three-term consonant cluster—for example in words like *screw*, *splint*, *skew*, and *square*. In other initial clusters these consonants must occupy the position immediately before the vowel and cannot have any other consonant intervening between them and the vowel.[3] Similarly, where they occur in final consonant clusters they must, again, occupy the place nearest the vowel, as in words like *melts* and *birds*.[4] The one exception to this is in words such as *snarl* or *world*, where /r/ comes between the vowel and /l/. We can, therefore, generalize the distribution of these sounds by saying that they must occupy a position in the syllable in immediate contact with the vowel (or with the vowel plus /r/), and that they are the only sounds permitted as the third term of a three-term initial consonant cluster.[5]

The fact that /w, j, r, l/ can be grouped together on articulatory and distributional grounds permits us to hope that they will have certain significant acoustic features in common. We should expect, then, that some of the acoustic cues which enable the listener to distinguish within the class will be found to lie on the same acoustic dimension, and, further, that these common acoustic features will serve, at least in part, to distinguish this class of phonemes from other classes.

METHOD

The inspection of acoustic spectrograms is a rewarding study which has already made a notable contribution to linguistic work, and it was upon such study that we relied for our preliminary observations. If a spectrogram has a fault, however, it is in the

[3] In many, perhaps most, dialects of English this would be an exclusive class-marker but, in the case of those dialects where "new" = /nu/, the phoneme /n/ shares this distributional characteristic.

[4] This sets them off from all English phonemes except /ŋ/ which must also occupy the immediately post-vocalic position, although it is quite unlike /w, j,r,l/ in its general distribution.

[5] For a more detailed statement of the combinatory latitudes of /w,j,r,l/ see J. D. O'Connor and J. L. M. Trim, Vowel, consonant, and syllable — a phonological definition, *Word*, 9.103-122 (1953).

abundance of its revelations, which lead an investigator to wonder what, in a given stretch of speech, is linguistically basic and what dispensable. Often, too, the acoustic features are only vaguely shown on a spectrogram, making it difficult to know the precise characteristics of those features which are revealed. These difficulties can to some extent be overcome by use of the pattern playback, a machine with which hand-painted patterns, resembling spectrograms more or less closely, may be converted to sound.[6] This provides a convenient method for making experimental modifications in various presumably important parts of the pattern and then evaluating the effects of these changes on the sound as heard. The present investigation was based primarily on this technique.

We were concerned in this experiment with the recognition of phoneme sequences as such. This is, of course, only one stage in the process by which speech is understood: the sounds perceived must be set against the pattern of both linguistic and extra-linguistic probabilities and a balance struck. In the experiments being reported here we have minimized the effect on intelligibility of any but the immediate sound contexts by working exclusively with nonsense syllables of the pattern CV. The vowels used were seven in number, corresponding approximately to the cardinal vowels [i, e, ɛ, a, ɔ, o, u].[7]

The experiments fall into two parts. In Part I we explored a wide variety of acoustic variations and attempted to determine, always by our own listening, which of these variations were important for the perception of /w, j, r, l/ and which were not. It is difficult to know precisely how many patterns and pattern changes were made and listened to, but the number is in the thousands. This is an important consideration only because it means that we were unable to control perfectly the contexts in which, and occasions on which, we listened to and judged the

[6] For a description of the pattern playback see F. S. Cooper, Spectrum analysis, *Journal of the Acoustical Society of America*, 22.761-762 (1950); F. S. Cooper, A. M. Liberman, J. M. Borst, The interconversion of audible and visible patterns as a basis for research in the perception of speech, *Proceedings of the National Academy of Sciences*, 37.318-325 (1951); F. S. Cooper, Some instrumental aids to research on speech, pp. 46-53 in *Report of the fourth annual round table meeting on linguistics and language teaching*, Washington, D. C. : Institute of Languages and Linguistics, Georgetown University (1953).

[7] See P. C. Delattre, A. M. Liberman, F. S. Cooper, Voyelles synthétiques à deux formantes et voyelles cardinales, *Le Maître phonétique*, 96.30-37 (1951).

various patterns. The possibility exists, therefore, that our judgments were made against a background or standard that varied somewhat from week to week and from one context to another. As a partial control against this possibility, and also in order to have judgments from a group of naive listeners, we carried out the experiment reported in Part II. For the purposes of the second part of the study we sampled from the various acoustic dimensions we had found to be important in Part I, arranged these stimuli in a random order, and presented them to naive college students for judgment as /w/, /j/, /r/, or /l/.

PART I

Figure 1 shows patterns that contain the cues we believe to be most important for the perception of /w, j, r, l/. These patterns are, in general, like those we have found in previous studies to be appropriate for other CV syllables in that they consist primarily of formant transitions, or frequency shifts, followed by a steady state.[8] It had been found in the earlier studies that the transitions contain important, and in many cases sufficient, cues for the perception of the consonants,[9] while, as is well known, vowel color depends largely on the frequency position of the steady-state formants.[10] As can be seen from these patterns, the distinctions among /w, j, r, l/ depend primarily on the transitions of the second and third formants. The second-formant transition is sufficient to distinguish /w/ from /r-l/ from /j/: for /w/ this transition originates

[8] We have drawn all our transitions as straight lines even though the transitions of real speech are necessarily curvilinear. Using straight lines enables us to control the patterns more precisely, and it does not significantly alter the auditory impression.

[9] See Liberman, Delattre, Cooper, Gerstman, *op. cit.*

[10] This has been demonstrated by several methods : from a statistical analysis of the vowel productions of many speakers by G. E. Peterson and H. L. Barney, Control methods used in a study of the vowels, *Journal of the Acoustical Society of America*, 24. 175-184 (1952); from tests constructed at the pattern playback by P. C. Delattre, A. M. Liberman, F. S. Cooper, L. J. Gerstman, An experimental study of the acoustic determinants of vowel color: observations on one- and two-formant vowels synthesized from spectrographic patterns, *Word*, 8.195-210 (1952); and from vowel synthesis employing electronic analogs of the vocal tract by several investigators. See: H. K. Dunn, The calculation of vowel resonances, and an electrical vocal tract, *Journal of the Acoustical Society of America*, 22.740-753 (1950); K. N. Stevens, S. Kasowski, C. G. M. Fant, An electrical analog of the vocal tract, *Journal of the Acoustical Society of America*, 25.734-742 (1953); K. N. Stevens and A. S. House, Development of a quantitative description of vowel articulation, *Journal of the Acoustical Society of America*, 27.484-493 (1955).

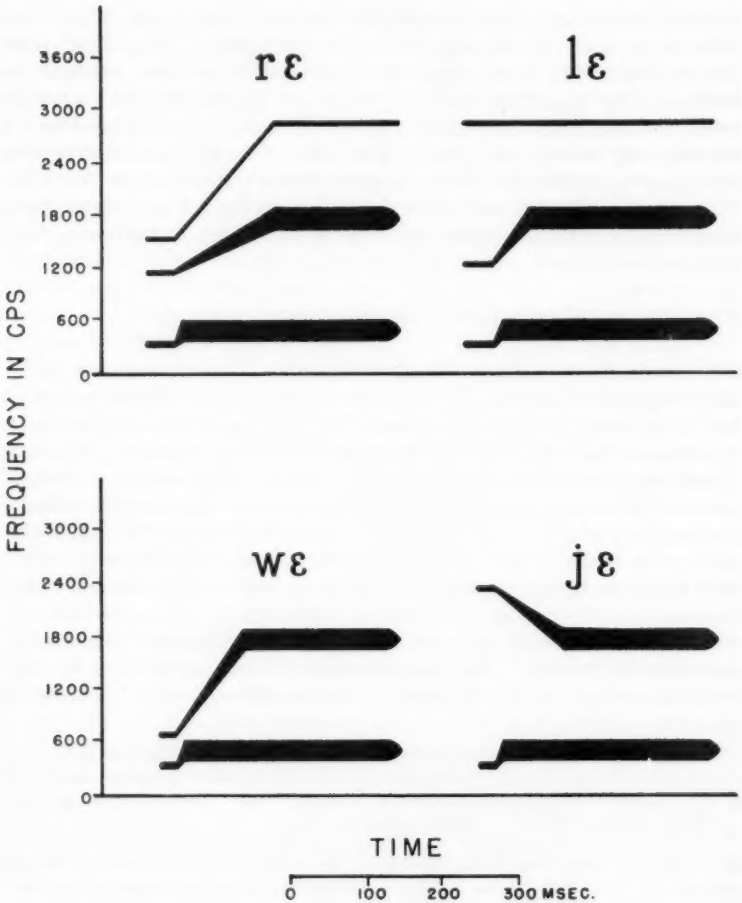


Fig. 1. Hand-painted spectrograms that produce reasonable approximations to /rɛ/, /lɛ/, /wɛ/, and /jɛ/ when converted into sound by the pattern playback.

at a low frequency, for /r/ and /l/ it begins nearer the middle of our frequency range, while /j/ requires that the transition start high. The distinction between /l/ and /r/ seems to depend primarily on the third-formant transition: for /r/ the third formant begins at a point just slightly above the second formant and then

rises to its steady-state level, while for /l/ the third formant starts at a level at least as high as the steady-state.

There is also at the beginning of each formant a relatively short steady state which, for convenience, we will refer to as a "steady-state onset." These steady-state onsets represent what we have previously referred to as the consonant loci,[11] and they are, quite obviously, the acoustic counterparts of the starting points of the consonant articulation. It is of considerable interest that the loci are explicit for /w, j, r, l/ because in this respect these four phonemes are different as a class from the stop and nasal consonants. In the case of the stops (and presumably for the nasal consonants, too) we had found for the second formant that the transition may not begin at the locus but must rather start at a point somewhat delayed in time. The steady-state onsets we see in Fig. 1 indicate that for /w, j, r, l/ the formants not only begin at the loci but spend from 30 to 50 msec. there before proceeding to the steady-state positions of the vowel.

Steady-state onsets similar to those of Fig. 1 were found in an earlier study to be necessary for the synthesis of the nasal consonants in initial position, but in this respect /w, j, r, l/ are very different from the nasal consonants in that the steady-state onsets are always continuous with the transitions for the former, but typically differ in frequency from the starting points of the transitions for the latter. This is so, presumably, because in the case of the nasal consonants the steady-state onset is produced as a consequence of the addition of the nasal resonator and is little affected by the (buccal) articulatory position, which nevertheless determines the starting frequency of the formant transitions. With /w, j, r, l/ the steady-state onsets reflect the start of the buccal articulation.

Having made these rather general comments about the acoustic cues that distinguish among /w, j, r, l/, and about certain characteristics of these four phonemes as a class, we should like to proceed to a more detailed discussion of our results.

Number of formants needed.

As shown in Fig. 1, a distinction may be made at the outset between /w, j/ on the one hand, and /r, l/ on the other, in that

[11] For a discussion of the "locus" concept, see P. C. Delattre, A. M. Liberman, F. S. Cooper, Acoustic loci and transitional cues for consonants, *Journal of the Acoustical Society of America*, 27.769-773 (1955).

the former can be synthesized satisfactorily with only two formants, while the latter require three. In the case of /w, j/ this is no more than might be expected, bearing in mind the well-known possibility of two-formant vowel synthesis, and the articulatory relationship between [i] and /j/ and between [u] and /w/. We find for /r/ and /l/, however, that a third formant is clearly necessary in order to synthesize these two consonants before the full range of vowels. As can be seen in Fig. 1, and as we will have occasion to discuss in greater detail later, the second-formant starting points are not very different for /r/ and /l/—they both begin in the middle of our frequency scale—so it is not surprising that we cannot distinguish these phonemes by means of two formants alone. This does not mean that the more-or-less common /r-l/ second formant will produce, with all vowels, a phoneme which is indifferently /r/ or /l/. In fact, we find with two-formant patterns that when the second-formant transition rises from the /r-l/ region to the frequency levels appropriate for front vowels we hear /r/, but when the transition falls to a back vowel we hear /l/.

Duration of steady-state onsets.

Steady-state onsets were found to be useful, not for any distinctive function they might have within our group of consonants, but to avoid a potential confusion with clusters of Stop+/w, j, r, l/. When the onset phase was omitted and the transition begun immediately, the resulting stimulus appeared to have an explosive beginning, a stop of some kind being clearly apprehended though not always identified with certainty. The duration of steady-state necessary to avoid this clustering effect differs as between /r, l/ and /w, j/. For /r/ and /l/ a steady-state onset period of about 50 or 60 msec. was needed—below this the explosive beginning started to make itself heard—whilst at greater durations, from 70 msec. upwards, the effect was of syllabicity, often leading to the identification of a schwa vowel before the consonant. For /j/ and /w/ the onset duration could apparently be smaller; 30 msec. was sufficient to eliminate the explosive beginning, and 40 msec. upwards gave the effect of a full vowel, /i/ or /u/, despite a reduction in intensity of the steady-state onsets *vis-à-vis* the following vowel formants.

Recent work suggests that the steady-state onsets can perhaps be dispensed with for /w/ and /j/ without causing an explosive beginning to the syllable, provided that the first-formant transition

begins at a sufficiently high frequency (discussed in the next section). We have also found it possible to hear /r/ without initial steady states, provided that the intensity onset is gradual, but we have never been able to synthesize an acceptable /l/ using transitions alone. We note that none of the four phones is harmed by the presence of appropriate steady-state onsets, and indeed, that spectrograms typically show, if not a steady state preceding the transitions, at least a very much slower rate of frequency change than pertains to the transition phase.

Starting frequencies of the transitions (steady-state onset frequencies).

In the preceding section we observed that appropriate steady-state onsets serve mainly to distinguish the class /w, j, r, l/ from other groups of phonemes. Since these onsets have durations up to 60 msec., it is not inconceivable that they might also themselves be identifiable as /w/, /j/, /r/, or /l/, but our observations suggest that for the most part they are not; rather, it is the transitions, particularly those of the second and third formants, which play the main part in distinguishing among the four phonemes. Steady-state /w/ and /j/ are quite out of the question since the absence of transitions can result only in /u/ and /i/. The onset of /l/ contains some identifying information, but naive listeners have considerable difficulty in perceiving the steady state as /l/. For /r/ there is more information in the steady-state onset, to the extent that the onset approximates an /r/-colored vowel.

Bearing in mind that the transition directions and extents are the most important cues for distinguishing among /w, j, r, l/, we will nevertheless specify these variables by reference to the frequency positions of the steady-state onsets, since they are always continuous in frequency with the start of the transition. These starting points provide a convenient way of describing the transitions because the onsets remain relatively fixed while transition directions and extents shift radically with changes in the formant positions of the following vowel.[12]

First-formant onset. It will be obvious from the patterns of Fig. 1 that the onset of the first formant does not distinguish among /w, j, r, l/ since, for the same following vowel, its form is unchanged throughout. It does, however, serve the very impor-

[12] Delattre, Liberman, Cooper, (1955), *op. cit.*

tant function of distinguishing /l/ from the nasal consonants. If the onset of the first-formant transition is located at 240 cps, the resulting sound will be identified more often as a nasal than as /l/ even when the second and third formants are given their best /l/ values, whereas with the first-formant onset at 360 cps, or above, the identification will be /l/. The best /l/ results were obtained with the steady-state onset located at or slightly below the frequency of the first formant of the following vowel, but no lower than 360 cps. In subsequent work the first-formant onset was placed at 480 cps for [a], and 360 cps for the remaining vowels.

For both /w/ and /j/ one would expect the frequency ot the first-formant onset to be in the neighborhood of 240 cps (the first-formant frequency of both /i/ and /u/), and this proved to be well founded. In fact, there was not much freedom to deviate from that value lest either of two complications occur: when the steady-state onset was located at 120 cps, the semivowels acquired stop-like beginnings, /bw/ or /gj/, an effect that we have discussed in other papers;[13] when the onset was placed much above 360 cps, the impression of /w/ was destroyed, while a non-speechlike whistle was introduced into /j/.

The influence of the first-formant starting frequency on /r/ was small. If a pattern contained second- and third-formant transitions satisfactory for /r/, the only difference made by a first-formant onset between 120 and 600 cps was a difference of "color", corresponding at the bottom of the range to extreme lip-rounding and at the top to lip-spreading.

We see, then, that /l/ imposes more stringent requirements upon the first-formant onset frequency than the other three phonemes. However, when we use values appropriate for /l/ we also satisfy the restrictions imposed by /w, j, r/.

Second-formant onset. The onset frequency of the second-formant transition has a very considerable effect on the perception of /w, j, r, l/. Other things being equal, high frequencies (about

[13] Delattre, Liberman, Cooper, (1955), *op. cit.;* A. M. Liberman, P. C. Delattre, L. J. Gerstman, F. S. Cooper, Tempo of frequency change as a cue for distinguishing classes of speech sounds, *Journal of Experimental Psychology*, 52.127-137 (1956). The first-formant locus for the stop consonants was found to be at the lowest frequency attainable with the pattern playback (120 cps). Any higher frequency caused the stops to sound less like stops and more like semivowels. This effect was slight, however, as compared to the strong influence of transition duration: patterns were heard as semivowels or stops depending on whether the transitions were more or less than 50 msec.

2760 cps) give /j/, low frequencies (about 600 cps) give /w/, and intermediate frequencies give either /l/ or /r/. The frequency ranges of /l/ and /r/ overlap.

For /w/ and /j/ the starting frequencies are, again as one would expect, in the neighborhood of the second formants of [u] and [i], respectively, but the opener and closer semi-vocalic allophones are clearly reflected in the differing ranges before different vowels. Before the vowel [i], the lower limit of the /j/ range has the same frequency as the second formant of the vowel (2760 cps), the glide effect being obtained by the transition of the first formant and the gradual increase in the intensity of the transitions; more satisfactory results are obtained, however, with a higher second-formant onset frequency and consequent falling transition. Before [e] (second formant at 2160 cps), the lower limit of the range is 2280 cps. One cannot start /j/ below 2280 cps for any following vowel without producing the effect of a glide from [ø] or [y]. There is virtually no upper limit for /j/, 3600 cps still being entirely satisfactory.

The frequency range of the second-formant onset for /w/ before [u] (second formant at 720 cps), is the very low and restricted one of 360-480 cps; before [o] on the other hand the upper limit is 600 cps, and for the remaining vowels, 840 cps.

For /r/ the frequency range of the second-formant onset before [i, e, ɛ] is 840-1560 cps, before [ɔ] 840-1200 cps, and before [o, u] 600-1200 cps. In the lower part of the range the acoustic counterpart of labio-velarization is heard, leading to confusion between /r/ and /w/ below the lower limit, whilst in the higher part of the range an effect of palatalization is heard.

For /l/ before [i, e, ɛ] the frequency range is 960-1800 cps, before [a] 840-1800 cps, and before [ɔ, o, u] it is 840-1680 cps. In the lower part of these ranges the /l/ was of a dark or velarized variety, the lower the frequency the darker the /l/; at the higher frequencies the /l/ was clearer or more palatalized, until, above the upper limits here given, the effect of laterality was lost, and was replaced by a vowel glide from [e] or from [ɪ].

We can therefore distinguish /w/, /l, r/, and /j/ by reference to the second formant. Although the starting point of the second-formant transition should be somewhat higher for /l/ than for /r/, this does not provide a reliable differentia between the two phones.

Third-formant onset. The third formant is the crucial factor, as we suggested earlier, in distinguishing /l/ from /r/, but contri-

butes little to the perception of /w/ and /j/. The third-formant starting frequency for /l/ is close to that of the vowel third formant, whereas the third-formant onset of /r/ needs to be lower in frequency, fairly close to the second-formant onset. This being the case, it is possible to pass from /l/ to /r/ by no other change in the pattern than a gradual lowering of the starting frequency of the third formant. If /w/ and /j/ are to have a third formant, the transition can be similar to that required for /l/, although as we have pointed out, such a third formant contributes very little to the perception of the semivowels.

We have already mentioned that the need for third-formant transitions in distinguishing /l/ from /r/ varies as between front and back vowels, so it should not be surprising to find these differences reflected in the permissible ranges of third-formant starting frequencies. For the vowels [i, e, ɛ], where the second-formant transition is rising and only two formants are required to hear /r/, the third-formant onset for /r/ can vary from 840 cps (the lower limit for the second-formant onset) to as high as 1920 cps; this means in effect that the frequency ranges of the second- and third-formant onsets are virtually identical for /r/, whereas the third-formant onset for /l/ must be no lower than the third formant of the vowel. For the vowels [ɔ, o, u], where the second-formant transition is falling and a two-formant pattern is heard as /l/, the restrictions are now reversed: the third-formant onset for /l/ may be as low as 1920 cps, but for /r/ it may be no higher than 1680 cps. It is possible, of course, to find starting frequencies for the second and third formants which cancel each other, so that neither /r/ nor /l/ is perceived with clarity; we may make /r/ and /l/ maximally discriminable, however, by placing the third-formant onset either as low as the second-formant onset permits, or as high as the third formant of the vowel.

Precise specification of formant starting frequencies. The reader will have noted that we summarized our observations about the starting frequencies of the second- and third-formant transitions in terms of ranges rather than as specific values. Thus, for example, we placed the second-formant onset frequency of /re/ between 840 and 1560 cps, but did not say that any particular value between these extremes gave a "better" /re/ than any other.

We believe that this is the most realistic way of stating our findings, since, in different idiolects and dialects, there are many phonetically discriminable sounds which are nevertheless readily identifiable as /r/. Similarly, two different frequencies within

our range will produce sounds which are distinguishable, but which are, in our estimation, both identifiable as /r/. We do not claim that every sound produced from different frequencies within a stated range makes an equally natural impression on us, or would do so on a wider selection of listeners—a very clear [l] initially may sound as odd to some listeners as a very dark [l] does to others—but we do believe that every such sound is recognizable phonemically.

The limits of the ranges are usually difficult to draw, since a given stimulus, evaluated phonetically, may contain auditory features reminiscent of both, say, /w/ and /r/, or /r/ and /l/; further, a stimulus identified by the phonetician as [ŏ] might well be categorized by the naive listener as /j/, for lack of any other oppositional category in which to place it, even though [ø̆] is not normally found as an allophone of /j/. We have always tried to draw the line at a point where the identification seems to us to be still positive; the same identification might be made beyond the limits we have laid down, but we would expect more confusion in these areas and more uniform responses within our ranges.

Our observations indicate that the range of starting frequencies is shifted somewhat (on the frequency scale) according to the following vowel, and this is no doubt a function of conditioned allophonic variation : we should suppose that to produce a given allophone we must change the starting frequency as we change the vowel. However, our frequency ranges reflect a good deal more than the allophones of a single idiolect or even of a single dialect, within which we would not necessarily expect so wide a range of frequencies.

Since our own judgments, whilst basically phonemic, are necessarily also partly phonetic, it will be particularly interesting to see, in Part II, whether the shifting of starting points in relation to the following vowel is confirmed by naive listeners judging on what we assume to be a purely phonemic basis.

Transition duration.

Transition duration is another factor in the perception of /w, j, r, l/. Its role seems to be similar to that of the steady-state onsets in that it does not serve to distinguish among the four phonemes (with the minor exception to be noted below), but does aid in differentiating /w, j, r, l/ as a class from other groups of consonants. If the transitions are of too brief duration, there is

confusion with nasals and stops, while with transitions that occupy too much time in moving from the steady-state onsets to the steady-state vowel formants, there is the danger of losing the consonant impression entirely in favor of a vowel of changing color. Between these extremes there is a middle range of durations which serves reasonably well for all four consonants, although the particular values that give the most realistic sounds are somewhat different for each phoneme.

A duration of 100 msec. is suitable for the second- and third-formant transitions of all four phonemes. Briefer values give slightly better /l/ identifications, while longer ones help /r/. When /l/ is made as brief as 30 msec., there is the possibility of confusion with the nasal consonants, but at 60-70 msec. this ambiguity has disappeared, leaving /l/ at its most generally satisfactory value. When the transitions of an otherwise satisfactory /r/ pattern are reduced to 50 msec. or less, one hears a retroflex flapped sound, but, on the other hand, durations of up to 300 msec. do not destroy the /r/ effect.

Experiments with two-formant /w/ and /j/ showed that durations of 50 and 100 msec. were satisfactory, though the latter gave a somewhat more realistic sound; below 50 msec. the effect was of a vowel plus stop or flap.[14] At 150 msec. /j/ was still good, but an otherwise satisfactory /w/ pattern, before the vowels [i, e, ɛ], was heard as /wr/; this effect is undoubtedly due to the large rising transition of the second formant coupled with a comparatively slow rate of frequency change. The rising second-formant transition is often sufficient to produce an /r/ without the parallel rising third-formant transition that we have found to be so potent a cue for /r/. We find that a relatively broad rising second formant produces nearly the same effect as second- and third-formant rising transitions running parallel and in near contiguity. As one would expect, the /wr/ effect disappears when the transition durations are made briefer (thus speeding up the rate of frequency change), or when a straight third formant is added.

The first-formant transition for /w, j, r/ may have a duration identical to those of the second and third formants. For /l/, however, there is something of a special problem, since even at the

[14] These effects correspond exactly to those obtained with naive subjects who were able to hear the series /bɛ, wɛ, uɛ/ or /gɛ, jɛ, iɛ/ solely as a function of changes in transition duration. See Liberman, Delattre, Gerstman, Cooper, *op. cit.*, and Footnote 13.

briefest durations used with second and third formants (50 msec.), the first-formant transition seemed to be too gradual, giving the impression of a vocalic glide. It is apparently not without significance that spectrograms of spoken /l/ show a very abrupt frequency change between the first-formant onset and the following vowel. We sought to obtain this effect synthetically by reducing the transition duration to 10 msec., with the result that /l/ identifications were definitely improved. We found, further, that the patterns for /w, j, r/ were not adversely affected by this special first-formant transition, and have incorporated it into all the patterns used in Part II.

PART II

In this section we shall describe the results we obtained when we presented some of the stimuli of Part I to phonetically naive listeners for judgment as /w, j, r, l/. We are concerned here only with the cues that distinguish among these four phonemes. Accordingly, we have varied only those aspects of the pattern that had been found in our earlier investigation (see Part I) to be important for those distinctions, and we have asked our listeners to restrict their choices to /w/, /j/, /r/, or /l/. In regard to those acoustic variables which distinguish these phonemes from other classes, we have selected values which favored /w, j, r, l/ and held them constant in all the stimulus patterns.

The patterns seen in Fig. 2 illustrate, for the vowel [e], the stimulus variables we used in this experiment. As shown in the top row, there were six second-formant transitions, so chosen as to cover the range that had been found earlier to be important for /w, j, r, l/. We should note, however, that the range is sampled in rather large steps. By converting these two-formant patterns into sound we produced six of our test stimuli. Another twelve stimuli were made by combining each of these various second-formant transitions with a straight third formant and with a rising third formant as shown in the middle and bottom rows of Fig. 2. The straight third formant and the rising third formant had seemed on the basis of our earlier research to be reasonably appropriate for /l/ and /r/, respectively. There was, then, a total of 18 stimulus patterns for the vowel [e]. In the bottom row of the figure we see that the rising third-formant transitions always started at a point just slightly above the starting point of the second-formant transitions.

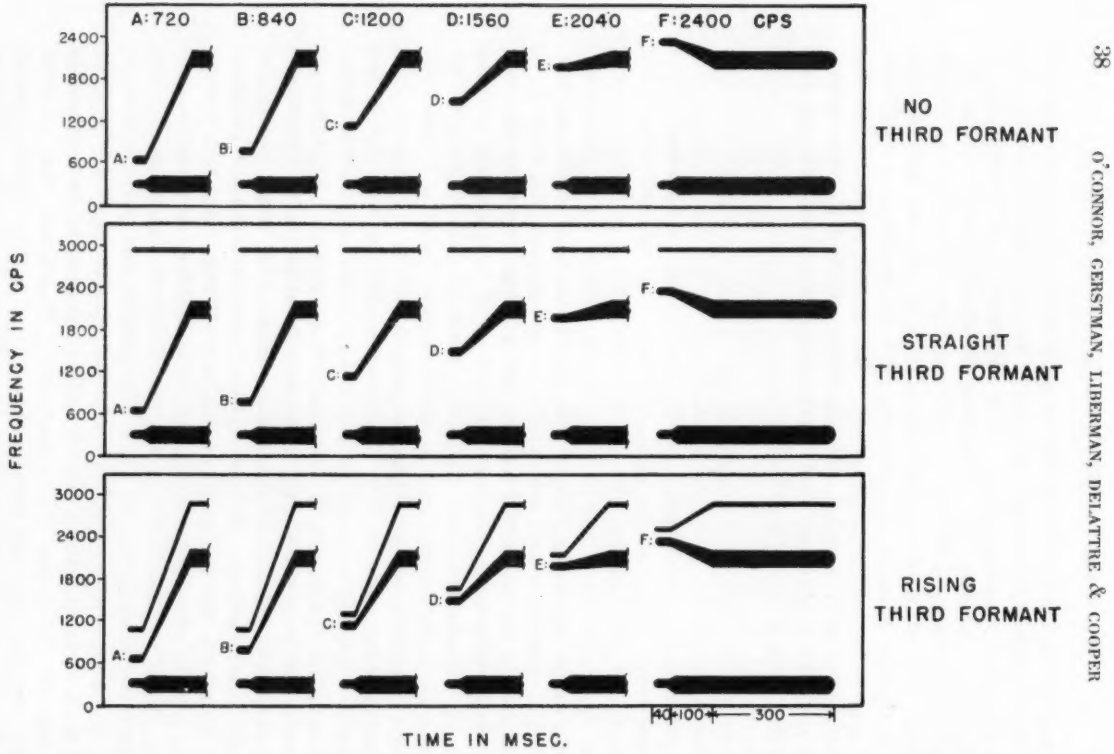


Fig. 2. Schematic representation of the stimulus variables used in Part II. The patterns shown are appropriate only for the vowel [e]. At the extreme right of each row are complete, hand-painted spectrograms illustrating some actual patterns from which the test stimuli were produced.

Fig. 2. Schematic representation of the stimulus variables used in Part II. The patterns shown are appropriate only for the vowel [e]. At the extreme right of each row are complete, hand-painted spectrograms illustrating some actual patterns from which the test stimuli were produced.

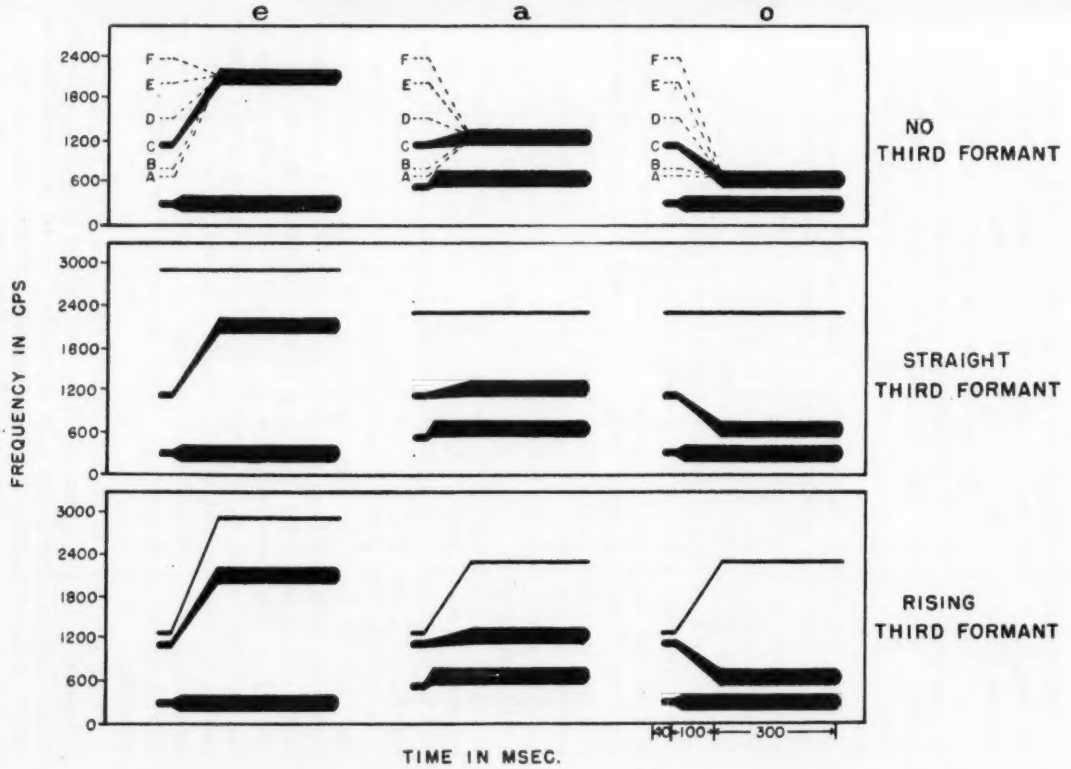


Fig. 3. Hand-painted spectrograms illustrating the stimulus variables of Part II with all three vowels [e, a, o].

There were, in addition, 18 stimulus patterns each for the vowels [a] and [o], constructed in essentially the same fashion as the stimuli for [e] in Fig. 2. Figure 3 summarizes the stimuli for [e, a, o] together. In the top row we see that the starting points of the second-formant transitions were identical for all three vowels. The middle and bottom rows illustrate, respectively, straight and rising third formants for a single second-formant starting point.

These patterns were converted into sound and presented in random order to a group of 44 phonetically naive listeners, all of whom were undergraduate students at the University of Connecticut. Each listener was asked to identify every sound as /w/, /j/, /r/, or /l/, and to guess if necessary.

The results of this experiment are to be found in the judgments of our listeners, all of which are shown in Fig. 4. There we have plotted the /w, j, r, l/ responses, expressed as per cents, against the second-formant onset frequencies, which are labelled on each abscissa by letter and identified by frequency in the key at the bottom of the figure. The three third-formant conditions (no third formant, straight third formant, and rising third formant) are the parameters, coded as shown in the upper right-hand corner. To avoid the difficulties in reading that would have been caused by many more or less overlapping curves, we have put the curves into separate plots according to responses (columns) and vowels (rows).

In general, the results obtained here are similar to those we described in Part I. The starting point of the second-formant transition distinguishes /w/, /r-l/, and /j/. This formant starts at a low frequency for /w/, at a middle frequency for /r/ and /l/, and at a high frequency for /j/. The third formant tends to distinguish /r/ from /l/, /r/ being helped by the rising third formant and /l/ by the straight third formant. Straight and rising third formants have little effect on /w/ and /j/, except that the rising third formant, in favoring /r/, tends to reduce the number of /w/ responses.

As we pointed out earlier, the stimuli that were presented for judgment in this experiment represent a rather large-step sampling of the various ranges, and it is, therefore, idle to draw conclusions which depend upon exact comparisons of the various curves, particularly in regard to the positions and heights of their peaks. However, some of these more specific effects are sufficiently large, and sufficiently consistent within the experiment, that we can

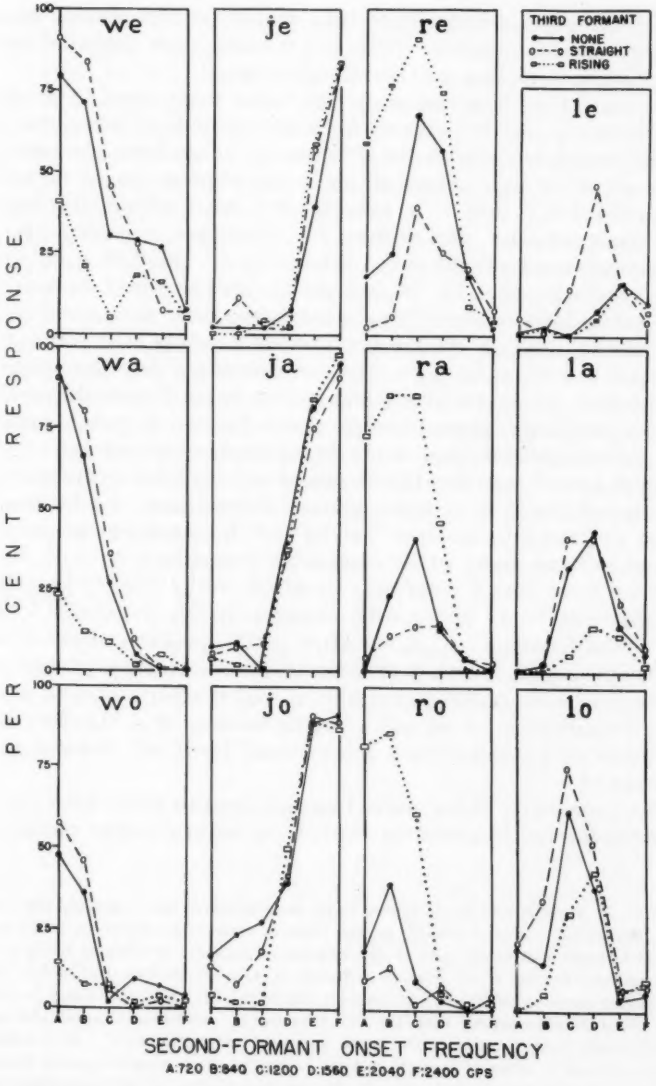


Fig. 4. Responses of 44 phonetically untrained listeners, plotted as a function of second-formant onset frequency for each of the three third-formant conditions.

be reasonably confident that they would be reproduced, at least in their grosser aspects, if the experiment were repeated with a much finer sampling of the stimulus range.

Perhaps the most important of these more specific results is the tendency of the peaks of both the /l/ and /r/ curves to occur at different positions on the abscissa as we go from the vowel [e] through [a] to [o]. These shifts in the response peaks mean that to produce the best /r/'s and /l/'s we must adjust the starting frequency of the second-formant transition according to the frequency level of the second formant of the following vowel. The starting point of the second-formant transition should be somewhat higher when the second formant of the vowel is high (as in [e]), and it should be somewhat lower when the vowel formant is low (as in [o]). We see, however, that the variations in the starting points of the transitions are quite small in relation to the frequency range through which the vowel formant moves. All of this confirms what we had previously observed and reported in Part I, and considerably increases our confidence in the result. The result itself is of some general importance, we believe, for when we take into account that /r/ and /l/ transitions begin at the actual loci (as many other consonant transitions do not), we see that we have here a clear case in which the frequency position of a consonant locus varies with changes in the frequency level of the vowel formant. This variation in the position of the locus is relatively small, which is fortunate for the utility of the locus concept as a simplifying assumption, but it clearly occurs, and we have remarked on it at some length because it is the first direct evidence we have had that a consonant locus will behave in this manner.[15]

The judgments of our naive listeners tend to agree with our own impressions (as reported in Part I) in several other particulars.

[15] K. N. Stevens and A. S. House have demonstrated that precisely this kind of locus movement can be expected, on the basis of their calculations, to occur for the second-formant loci of /b/ and of /g/: Studies of formant transitions using a vocal tract analog, *Journal of the Acoustical Society of America*, 28.578-585 (1956). In our own earlier research on the second-formant loci of /b/, /d/, and /g/, we had to contend with the difficulties arising from the fact that the second-formant transitions do not (and, indeed, cannot) begin at the locus, but must only point to it. As a result, we had to rely on a series of straight second formants (at various frequency levels) in order to find the /b/, /d/, and /g/ loci. By the very nature of that procedure, it was impossible to detect the kind of locus movement that we have here found with /r/ and /l/ and that Stevens and House would expect to find with /b/ and /g/.

We see in Fig. 4, for example, that the best starting point for the second-formant transition is at a somewhat higher frequency for /l/ than for /r/, though this difference would not appear from these data to provide a very good basis, in itself, for the perceived distinction between these two phonemes. It is also reasonably clear from the results with naive listeners that with two formants alone, /l/ is relatively good with the back vowels, while /r/ is relatively good with the front vowels.

It had been our feeling in working with these sounds that our /l/ was, at best, inferior to the other three phonemes, but we were somewhat surprised to discover how very poor our naive listeners found it. We believe that certain rather detailed changes in the first-formant steady-state onset and transition will considerably improve the /l/ for our naive listeners.

We see in Fig. 4 that /w/ before the vowel [o] was rather poor. This is not surprising when we consider that there were no patterns for [o] in which the starting point of the second-formant transition was below the steady state of the vowel (720 cps). Starting the transition at a somewhat lower frequency will probably improve the /wo/ considerably.

Haskins Laboratories.
New York, N. Y.

ON DISCRETENES AND CONTINUITY IN STRUCTURAL DIALECTOLOGY*

EDWARD STANKIEWICZ

In a recent article, "Is a Structural Dialectology Possible?",[1] Uriel Weinreich raised a number of problems which had figured in only a general and programmatic form in the writings of the pioneers of structural dialectology.[2] Although the actual achievement of some of these linguists gives the question a rather rhetorical cast,[3] Weinreich's article has the merit of indicating the points of divergence and overlap between traditional and structural dialectology and of stating again the importance of dialectology for structural linguistics. However, some of the issues which are of interest to the dialectologist as well as to the general linguist warrant further discussion and clarification. This article proposes to take up and to elaborate some of the suggested approaches and the questions which were left open in Weinreich's study.

The central argument of this paper concerns the selection of phonemic criteria as the basis for a structural description of dialects. The illustrative material is drawn from works in Slavic dialectology, among which the number of phonemic studies is very scarce.[4] In the present state of dialectology, the structural

* A shorter version of this paper, under the title "The Selection of Structural Criteria in the Grouping of Dialects," was read at the Thirty-first Annual Meeting of the Linguistic Society of America, December 28, 1956.

[1] *Word* 10.388-400 (1954).

[2] Especially in the article of N. S. Trubetzkoy, « Phonologie et géographie linguistique », *TCLP* 4.228-34 (1931).

[3] Among the earlier studies in structural dialectology are B. Havránek's "Nařečí česká", *Československá Vlastivěda* III, Prague, 1934, pp. 84-218, and G. Laziczius' *Bevezelés a fonológiába*, Budapest, 1932, pp. 62-74.

[4] To these belong the quoted study of Havránek, the less consistently phonemic description of the Russian dialects in R. J. Avanesov's *Očerki russkoj dialektologii* I, Moscow, 1949, and the author's "The Phonemic Patterns of the Polish Dialects," *For Roman Jakobson*, The Hague, 1956, pp. 518-531. The other main works referred

linguist is compelled to rely upon historical or phonetic descriptions which he has to interpret phonemically. The results of such a task can therefore be stated only tentatively. They are offered here primarily as an exemplification of the general problems and of a methodology which could be applied in other areas.

1. The classification of dialects on the basis of their phonemic inventories should be considered as the first step towards an exhaustive description of linguistic areas. The study of phonemic distribution and of other levels of analysis can and should complement the classificatory criteria discussed below. The various types of "akan'e" found in Russian and Byelorussian areas or the distribution of voiced and voiceless consonants in the Ukrainian dialects provide interesting distributional criteria for a supplementary classification of these territories. The "mixed" dialects in the border areas of Poland are excellent examples of speech communities which share their phonemic systems with one area and their grammatical systems and vocabularies with another.[5] Any attempt to fit these dialects into one major type or language would be artificial. However, a classification based on phonemic criteria is the most likely to be successful and relevant. Traditional dialectology, with all its shortcomings, has been almost entirely constructed on phonetic principles, which is suggestive of the fact that dialects are most clearly differentiated phonemically. The phonetic materials which have been accumulated by generations of scholars, some of which remained unaffected by the all-pervasive passion for historical reconstructions,[6] are therefore easily accessible to the linguist, whereas few data have been collected pertaining to grammar and syntax. Where such data exist, they indicate

---

to deal with the Western and Eastern Slavic dialects and are as follows: K. Nitsch, "Dialekty języka polskiego," in "Język polski i jego historia," *Encyklopedia Polska*, 3.2. pp. 238-343 (1915); S. Urbańczyk, *Zarys dialektologii polskiej*, Warsaw, 1953; N. N. Durnovo, N. N. Sokolov, D. N. Ušakov, *Opyt dialektologičeskoj karty russkogo jazyka v Evrope*, Moscow, 1915; J. Ziłyński, "Opis fonetyczny języka ukraińskiego," *Prace Komisji Językowej, PAU*, nº 19 (1932); F. T. Žilko, *Narysy z dialektologiji ukrajins'koji movy*, Kyjiv, 1955; B. F. Karskij, *Belorusy*, II, 1, Moscow ₁1955; V. Važný, "Nařečí Slovenská," *Československá Vlastivěda* III, Prague, 1934, pp. 219-310.

[5] For a survey of these dialects with numerous examples of phonemic and morphological diffusion, see Z. Stieber, "Sposoby powstawania słowiańskich gwar przejściowych," *Prace Polskiej Akademii Umiejętności*, No. 27 (1938).

[6] One such outstanding dialectologist who ignored historical problems was O. Broch, whose book *Slavische Phonetik* (1911) is still unsurpassed in its scope; also his pupil, J. Ziłyński.

relative uniformity over wide areas.[7] Lexical differences, which are the most striking to the layman, cannot be fruitfully applied in a structural approach; they reflect ethnic, social and geographic conditions of different areas and lend themselves least to systematization. Being of a non-systemic and cultural character, lexical criteria are otherwise utilized to define social dialects (argot, slang, professional languages). Phonology, which is the most advanced branch of modern linguistics, suggests itself, finally, as the most logical one within which we can determine criteria for the grouping of dialects.

2. Some linguists are of the opinion that dialect studies are incomplete or even impossible if they do not make use of extra-linguistic criteria. "The concept of language area *(Sprachlandschaft)* has practically replaced that of dialect as the central interest in most geographic work... Tests of mutual intelligibility, dialect sociology and statistical correlation methods...yield an insight into the makeup of a continuously varying language area which supplements, if it does not supersede, the results derived by other methods."[8] While one can agree that the results obtained by other methods and in other fields can ultimately be correlated with the results yielded in a linguistic dialectology, they must be considered of secondary importance to the linguist. The use of extra-linguistic criteria, which are by no means better definable than the linguistic criteria, introduces new variables in the study of dialects which are likely to obscure and to conflict with the results obtained by linguistic methods.

The concept of "Sprachlandschaft" as defined by German dialectologists was intended to overcome, or rather to by-pass the difficulties which arose in traditional dialectology (see § 4). By substituting for the notion of discrete dialects the concept of "dynamic" culture-zones ("geistiges Kraftfeld"), and by declaring the whole problem of dialect boundaries to be a "static, rigid concept," the historical dialectologist shifted the emphasis to the study of geographical areas, as networks of cultural, political, economic and linguistic change. The deterministic question of "causes" remains at the center of this approach.[9]

[7] Cf. S. Urbańczyk *(op. cit.)*, pp. 31-45, J. Avanesov and V. G. Orlova, "Voprosy izučenija dialektov jazykov SSSR," *Voprosy Jazykoznanija*, 1953, 5, p. 42. B. Havránek *(op. cit.)*, pp. 96-100.

[8] U. Weinreich *(op. cit.)*, pp. 397, 398.

[9] The quotations are drawn from A. Bach, *Deutsche Mundartforschung*, Heidelberg, ²1950; for the concept of "Sprachlandschaft", cf. especially pp. 56, 76, 122.

Intelligibility tests likewise have many shortcomings as a measure of dialect unity and differentiation. The ability to understand speakers of different speech communities depends to a large extent on training in code-switching and on individual talent, and it correlates more closely with lexical similarity than with structural correspondences between linguistic systems. Slovaks, according to Stieber,[10] are better understood by Russians and Serbians than by Czechs, although the phonemic and grammatical patterns of Slovak and Czech differ less than the structures of Slovak and Russian or Serbo-Croatian.

The reliance of some linguists on the opinions of native speakers in measuring differences and similarities between dialects is likewise deceptive. Social and cultural attitudes color the evaluation of one's own or foreign speech habits. This often accounts for linguistic traditionalism or for the ready acceptance of other linguistic patterns. The peasants as well as the intelligentsia of Silesia consider the speech of other Poles "foreign" ("ŕǫzǫ po polsku"), whereas in other areas of Southern Poland, even peasants try to imitate literary Polish and do not recognize the differences between their local speech and the standard language.[11]

A structural dialectology can successfully supersede traditional dialect studies by relying primarily on internal, linguistic standards.

3. In modern linguistics there is, as Weinreich puts it, "an abyss between structural and dialectological studies." The situation in dialectology is very much like the one which prevailed in historical linguistics, where standards differed from those applied in synchronic descriptions.

The central problem in dialectology is the problem of grouping divergent systems, distributed in geographical proximity, within a higher framework of similarities or of convergence. Dialects are thus viewed as varieties of what is in the layman's and in traditional terminology a "language" and of what has been recently referred to as an "overall pattern" or "diasystem."[12]

[10] Z. Stieber, ' Stanowisko mowy Słowaków," *Prace Filologiczne*, No. 17, pp. 37 ff. (1937).

[11] K. Nitsch, *Wybór pism polonistycznych* I, Wrocław, 1954, pp. 216-217. For the role of prestige factors in the identification of dialects, see also N. v. Wijk, *Les langues slaves, de l'unité à la pluralité*, The Hague, 1956, pp. 49-50. The rudimentary state of other extralinguistic procedures is discussed by Weinreich himself; *op. cit.*, p. 398.

[12] U. Weinreich *(op. cit.)*, p. 390.

It is obvious that this supersystem, which shares features common to all dialects, is different from the standard language (or languages), which is sometimes based on one prestige-dialect or which is a culturally determined compromise formula of various dialects; nor can it be defined in terms of extra-linguistic (e.g. political, administrative or cultural) boundaries.

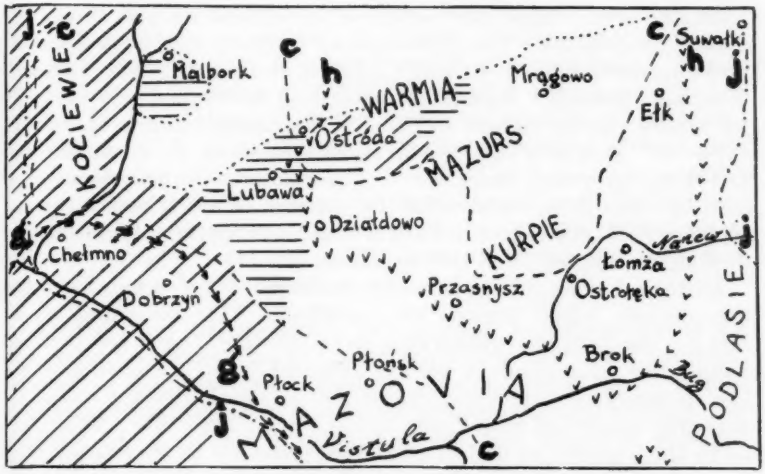
In comparing dialects which are similar enough to permit mutual intelligibility between speakers of neighboring areas, but different enough to be identified as such by most speakers, dialectology is confronted with the problem of defining the relation between continuity and discreteness. If it is not to dissolve into a typology which compares systems or sub-systems of most remote or highly differentiated languages simply in terms of difference or identity, dialectology has the task of stating the validity of the boundaries which it sets up and of calibrating the degree of similarity between the different regional varieties.

4. Operating with different criteria of what constitutes similarity and difference, structural dialectology is bound to face different problems and to get different results from traditional dialectology.

Traditional dialectology has been mainly historical and atomistic. It selected certain features of an older historical stage and mapped the territorial distribution of their etymological reflexes. To paraphrase Trubetzkoy, the pages of dialect studies were spangled with stars of reconstructed forms. In this approach the unity of different dialects was never questioned, since it was implicit in the method: it coincided with the reconstructed historical unity. Within the neogrammarian concept of a series of historical "splits," dialects were viewed as mere branches of an originally unified protolanguage.

The existence of a standard language, of political boundaries, and of other external criteria has, however, played an important and often decisive role in assigning the speech of certain areas to one or another "language."[13] Although dialectologists have generally tended to consider older features as more important than those of a more recent development, there has been no uniformity in the selection of their classificatory criteria; they varied from language to language and often from researcher to researcher.

[13] The claims of Serbian and of Bulgarian linguists to some Southern Slavic dialects on the basis of "typical" Serbian or Bulgarian features is a case in point. See S. Mladenov, *Geschichte der bulgarischen Sprache*, Berlin & Leipzig, 1929, p. 340.



Map. 1. Phonetic boundaries of Northeastern Poland
(Based on S. Urbańczyk, "Zarys dialektologii polskiej")

The recognized classification of the Polish dialects has been based on the different reflexes of the hushing spirants and affricates *š, *ž, *č, *ǯ and their priority over other features is indicated on the map (cf. map 1) by the shaded and blank areas. In the blank areas, *š, *ž, *č, *ǯ coalesced with *s*, *z*, *c*, *ʒ* (the so-called "mazurzenie" dialects); in the areas marked with horizontal lines, *š, *ž, *č, *ǯ changed into *ś*, *ź*, *ć*, *ʑ́* or fused into a single series *ś́*, *ź́*, *ć́*, *ʑ́* (the so-called "siakanie" dialects); whereas in the areas marked with slanted lines, the original three series *š*, *ž*, *č*, *ǯ*; *ś*, *ź*, *ć*, *ʑ́*; *s*, *z*, *c*, *ʒ* have remained intact. The "less important" features have served as the basis for further subdivisions and are indicated on the maps by lines. The area in which *å changed to *a* is delimited on the map by line (c); (g) indicates the area in which voiced consonants occur before vowels and sonorants in final position; (h) marks the southernmost limit of the pronunciation *ẏara* (Standard Polish /v'ara/); (j) delimits the area with the pronunciation *śvał* (Standard Polish /śf'at/). The above classification, which was established by K. Nitsch,[14] has served as the basis for all modern Polish dialect studies. In other Slavic

[14] In "Próba ugrupowania gwar polskich," *Rozprawy Polskiej Akademi i Umiejętności, Wydział Filologiczny*, Ser. I, I, pp. 336 ff. (1910).

areas, however, there has been less agreement on the choice of dialect boundaries. The Byelorussian territory has been divided into two main dialects, by some scholars on the basis of a phonetic principle (Karskij's bipartition according to the reflexes of **r'*), by others chiefly on the basis of a distributional principle (the types of "akan'e"). A classification into three dialectal groups has been proposed recently.[15] The relatively small Slovenian territory has been subdivided by various dialectologists, from Kopitar to Ramovš, on the basis of different criteria, into a number of dialects ranging from two to about fifty.[16]

Historical dialectologists were sometimes well aware of the arbitrariness which marred their method. This awareness gave rise to the famous controversy over whether there are dialect boundaries. Although it partly stemmed from the challenge of the neogrammarian "blind regularity of sound-changes" and is generally connected with the name of H. Schuchardt, the scepticism as to the validity of dialect boundaries was by no means limited to anti-neogrammarian quarters.[17] The controversy (between G. I. Ascoli and H. Schuchardt, P. Meyer and G. Paris) brought out the essential difficulties inherent in the methods of historical dialectology. Whereas Ascoli maintained that a "particular combination" of sounds, i.e. a bundle of isoglosses, is a sufficient basis for setting up dialect boundaries, his opponents argued that the lack of objective, general criteria in the selection of features predetermines the arbitrariness of all dialect divisions. The demarcation lines of dialect maps do not overlap, because they reflect isolated linguistic facts which have their own history; any division is thus an artificial construct, a "definitio nominis," and not a "definitio rei." The assumption of discrete dialects should be considered no more than a convention, whereas only the linguistic continuum is real.[18] The "fictional" view, that "in reality there are no dialects" (as was stated by G. Paris),[19]

[15] N. T.Vojtovič, "O dialektnoj osnove belorusskogo literaturnogo jazyka", *Voprosy Jazykoznanija*, 1954, pp. 26-41.

[16] F. Ramovš, *Dialektološka karta slovenskega jezika*, Ljubljana, 1931; *idem*, *Kratka zgodovina slovenskega jezika*, Ljubljana, 1936, pp. 107-113.

[17] This scepticism is clearly expressed in H. Paul's *Prinzipien der Sprachgeschichte*, Halle, 1880, pp. 237-240.

[18] A lucid account of the controversy is found in the article by L. Gauchat, "Gibt es Mundartgrenzen ?" *Archiv für das Studium der neueren Sprachen und Literaturen*, 111.365-403 (1903).

[19] In « Les parlers de la France », *Leclure faite à la réunion des Sociétés savantes*, 1888, reprinted in his *Mélanges linguistiques* II, Paris, 1907, pp. 432-448.

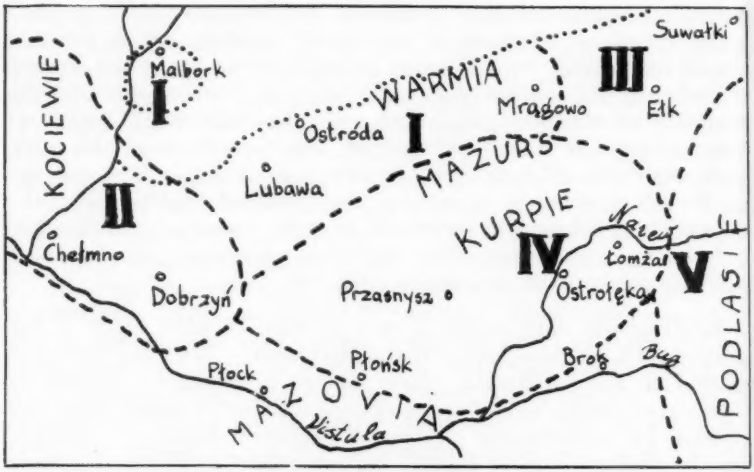
prevailed in dialectology. This view led not only to a phonetically complete atomization in dialect studies, which concentrated on tracing the history of isolated words (as in the French dialectological tradition of J. Gilliéron), but it was adopted by the majority of linguistis, including some forerunners of structural linguistics.[20] It is interesting that the champions of this view took the same attitude towards historical change. The concept of discrete systems in space was thus rejected together with the notion of discreteness in time, and linguistic "reality" was ascribed only to temporal continuity. G. Paris expressed this view in a succinct phrase : "Nous parlons latin."[21]

5. Structural dialectology is in a position to solve many of the problems which beset traditional dialectology. With the emphasis shifted to the synchronic plane and by selecting its criteria in terms of relevance and of systemic organization, it gained objective criteria for determining differences between linguistic systems. On the phonemic level, dialects are recognized as different if their phonemic inventories differ from each other. The phonemic maps drawn by the structural dialectologist will of course bear little resemblance to the traditional maps. Map 2 shows the territorial distribution of phonemic patterns in Northeastern Poland.[22] Area I indicates the dialects which have seven vowels and the minimal number of consonants (23 or 24) found on Polish territory. Area II has a higher roster of consonantal phonemes (28), but shows a generally lower number of vowels (5). Area III admits a number of consonantal phonemes equal to those in Area II, but they participate in different oppositions; namely here we find the opposition between hard and soft velars, which is absent in the other areas on the map, but not the opposition *š, č, ž́/ś, ć, ź́* (strident vs. mellow), which is found in Area II. In Area IV we find generally 25 or 27 consonants, and the highest number of vocalic phonemes, including nasal vowels (10). Area V

[20] See A. Meillet, *Les dialectes indo-européens*, Paris, ₄1950, p. 3, and F. de Saussure, *Cours de linguistique générale*, Paris, ₄1949, p. 279. In de Saussure we find, however, also some statements concerning the possibility of viewing dialects as discrete systems; cf. p. 276.

[21] This view too was embraced by the leading linguists before the rise of modern diachronic linguistics; see A. Meillet, *Linguistique historique et linguistique générale*, II, Paris, 1948, p. 81 ; F. de Saussure *(op. cit.)*, p. 122.

[22] This map shows only the approximate boundaries of the phonemic systems in Northeastern Poland. For a more accurate description and delimitation of the dialectal areas, see "The phonemic patterns of the Polish dialects" *(op. cit.)*.



Map. 2. Phonemic areas of Northeastern Poland

has seven vowels and a high number of consonants (30), including the opposition between hard and soft labial consonants, which is absent in the other areas. A comparison of the two maps reveals that the phonemic map ignores certain features which figured on map 1 as belonging to another level of description. For instance, the historically "important" distinction between the "mazurzenie" and the "siakanie" dialects is of secondary importance to the structuralist, inasmuch as both types have basically two series of spirants and affricates (with the exception, that is, of /ž/, of a different origin, which is opposed to /ź/ only in the "mazurzenie" dialects), instead of the three series which are found elsewhere. The different "realization" of the hushing consonants in both types is a redundant, not a distinctive phenomenon. The isoglosses (g) and (j) are distributional and lexical. Of "primary" importance to the structuralist are lines (c) and (h), which point to phonemic distinctions. As opposed to the traditional map, with its "primary" and "secondary", i.e. phonetic and distributional features, the structural map is marked by a homogeneity of features and by clearly defined areas.

6. If the main difficulty of historical dialectology lay in formulating discreteness, structural dialectology is faced with the

task of defining continuity and similarity between discrete local systems, and of grouping them into higher types.

In search of an "overall pattern," some linguists resorted to a graphic device in which "cover symbols" stand for various phonemes or phonemic sequences.[23] It is obvious that loosely applied transcriptional symbols, which overdifferentiate or underdifferentiate phonemic distinctions, can be stretched to cover not only the dialects of "American English" or of "Russian," but a variety of highly differentiated languages. Such a pattern is not an abstraction of features which are constant in a variety of closely related dialects, but is a fictitious construct.[24]

An extreme and purely formal structuralism may question the possibility of comparing parts of systems on the grounds that patterns are closed and unique configurations. However, few linguists would in practice endorse such a radical position. The historical acquisition of the phonemes /f, f'/ has produced no other change in the phonemic pattern of Standard Polish than the filling in of an empty slot within the framework of existing phonemic correlations. Similarly, Polish dialects in which there is the opposition between /ž/ and /š/ can be described in terms of the same distinctive features as those which lack this opposition. Under the influence of the literary language and of neighboring dialects, the voiceless spirant /š/ easily spreads into the dialects in which it is not phonemic, whereas dialect speakers preserve the opposition between the unrounded /a/ and the rounded /å/, even when they try to speak the literary language, in which this opposition is absent.[25] It is, therefore, also insufficient to compare systems simply in terms of their inventories; such comparisons do not amount to more than listing.

The classification of Slavic languages into phonological types, proposed by Isačenko,[26] which is based on a quantitative compar-

[23] G. L. Trager and H. L. Smith, Jr., *An Outline of English Structure* (= *Studies in Linguistics, Occasional Papers 3)*, Norman (Okla.), 1951, esp. pp. 20, 35.

[24] Cf. also the pertinent remarks of U. Weinreich *(op. cit.)*, p. 395.

[25] This fact has been recorded not only by linguists, but also by Polish novelists. In a novel by E. Zegadłowicz *(Zmory)*, the peasants do not make the distinction between /š/ and /ž/, but differentiate /a/ from /å/; high-school students coming into contact with the literary language, adopt the distinction between /š/ and /ž/, but preserve the distinction between /a/ and /å/, whereas the intelligentsia has taken on the literary language. For a different, sociological interpretation of this phenomenon, cf. Z. Stieber, "Sposoby powstawania słowiańskich gwar przejściowych" *(op. cit)*., p. 14.

[26] A. V. Isačenko, "Versuch einer Typologie der slavischen Sprachen," *Lingua Slovaca*, I.64 ff. (1939); cf. also R. Jakobson, *Slavic Languages*, New York, 1955, p. 16.

ison of the vocalic and consonantal sub-systems, has the disadvantage of confusing these types with the literary languages and of ignoring the variety of phonemic inventories found in the local dialects. It is this variety that makes it impossible to speak of phonemic "supersystems" or phonological types as if they were actual languages defined by political boundaries or by a common standard language; within the variety of dialects found in some areas we can, however, abstract a number of features, which sets it apart from varieties distributed over other geographical regions.

7. One of the measures of similarity between local systems with partially different phonemic inventories lies in their utilization of the same ultimate phonemic components. For instance, all Polish dialects share the distinctive features of consonantality, continuity, compactness, gravity, voicing, nasality and stridency, but none of them utilize prosodic features. On the contrary, the dialects east of the Polish linguistic territory (Russian, Ukrainian, Byelorussian) have the distinctive feature of stress, and the dialects south of Poland (Czech, Moravian, Slovak) have the distinctive feature of length. The Moravian (Lach) and Slovak (Čadce and Eastern Slovak) dialects which lack phonemic length, and the Ukrainian (Lemki) dialects which lack phonemic stress, form separate idalectal zones which share a number of phonemic features with the neighboring Polish dialects,[27] although distributionally and morphologically they have most features in common with the Moravian, Slovak and Ukrainian dialects, respectively. Other oppositions, such as tense vs. lax and flat vs. plain, are territorially more restricted than the prosodic features. They can provide the basis for further subdivisions of the Polish, Russian or Czech areas.

Comparing, furthermore, the common core of phonemes which are found in the dialects of wider linguistic areas, we notice that they share certain phonemes which are absent in other linguistic areas. For example, most Russian, Polish and Ukrainian dialects have in common the distinctive feature of palatalization, which sets them apart from other Slavic dialects, but the palatalized

[27] Among these features are penultimate stress, palatals in place of palatalized dental consonants, one phoneme /x/, instead of /h/ and /x/, the sporadic appearance of soft labial consonants. For more details, cf. N. v. Wijk, "Die Tschechisch-polnischen Übergangsdialekte," *Medellingen der Koninklijke akademie van vellenschappen* LXV, Ser. A, 1928, pp. 1-36; B. Havránek *(op. cit.)*, esp. pp. 184, 188; J. Ziłyński *(op. cit.)*, esp. pp. 85-90, 155-165.

consonants found in the respective dialects belong to different series. In most Russian dialects, all consonants, except the palatals, have palatalized counterparts; in the Polish dialects only hard velars and labials admit palatalized counterparts, and in the Ukrainian dialects only hard dentals and liquids admit palatalized partners. Nasal phonemes are found in all Slavic dialects; but in the Polish dialects we also find nasal vowels. The feature of length is distinctive in most Czech and Slovak dialects; however, in the Czech dialects proper, the short vowels /e, o/ have no long equivalents, whereas in the Slovak dialects the short and long vowels form symmetrical patterns.[28]

We can thus conclude that in areas which show the same distinctive features, their combinatory possibilities may be quite different. Re-stating the above examples, we would say that in the Russian dialects, sharping is incompatible with the consonantal compact-acute features; in Polish it is incompatible with the consonantal acute features; in Ukrainian it is incompatible with the consonantal compact and diffuse-grave features. The feature of nasality can combine in the Polish dialects both with consonantal and vocalic features, whereas in other Slavic areas it is combined only with the consonantal feature.[29] The feature of length is superposed in the Czech dialects only on compact and diffuse vowels, whereas in the Slovak dialects the combination of vocalic features with length is less restricted.

Another criterion of dialectal homogeneity is gained if we give due recognition to the coexistence of different styles within phonemic systems. As structural transformations in time are attenuated by the existence of certain phonemes which are unstable in the pattern,[30] variation in space is graduated by zones of transition in which some phonemic distinctions are not obligatory but optional, and as such endowed with social connotation. In interdialectal communication, speakers may resort to the phonemes which are "potential" in their system, but obligatory in the system of their interlocutors. As it is, these alternants are often known to them from the speech of older members of their own

[28] Cf. B. Havránek *(op. cit.)*, p. 103; J. Orlovský and L. Arany, *Gramatyka jazyka slovenského*, Bratislava, 1947, pp. 35-50.

[29] Nasal vowels are also found in some areas of Slovenia; cf. L. Tesnière, "Les voyelles nasales slaves et le parler Slovene de Replje," *Slavistična Revija* III, 3-4, pp. 263-265 (1950).

[30] B. Malmberg, "Bemerkungen zum quantitativen Vokalsystem im modernen Französisch" *Acta Linguistica* III, 1, pp. 44 ff. (1942).

speech community. Thus they find projected in space the facts which pertain to the dimension of time. For instance, the opposition between nasal and oral vowels, which is obligatory in some Polish dialects, is in neighboring dialects on the periphery of the linguistic norm: it is used by the older generation and it may appear in certain contexts and in the careful speech of younger speakers.[31] Similar is the status of the phoneme /r'/ in the "transitional" Ukrainian dialects, and of /c/ or /č/ in the Central Russian dialects.[32]

The admission of a broader frame of phonemic relevance enables us likewise to discern continuity, where on the basis of strict phonemic criteria we would see only discreteness. The phonetics-phonemics dualism has been rejected by a number of linguists on the grounds that the features endowed with other than distinctive functions are also part of the linguistic code and hence not irrelevant.[33] The distinctive features, being the least dispensable in the functioning of a sound system, are most readily identified by the speakers; "it would be deceptive, however, to believe that they are trained to ignore all the rest in speech-sounds."[34] Trubetzkoy suggested that the economy or a system can be evaluated in terms of the utilization of its features for various functions, i.e. the distinctive, configurational ("delimitative") and redundant ("associative").[35] For example, in Standard Russian we find that the feature of gravity is distinctive in the consonants, but redundant in the vowels; in Standard Polish the feature of sharping is distinctive in the velar stops /k', g'/, but configurational in the velar spirant [x']. In Standard Czech, in which the feature of sharping is not distinctive, it has neither a redundant nor a configurational function.

The study of the non-distinctive features, especially with relation to the distinctive features, reveals multiple points of structural

[31] W. Doroszewski, *Studia fonetyczne z kilku wsi mazowieckich*, Wrocław, 1955, esp. pp. 19, 38, 48.

[32] J. Ziłynski *(op. cit.)*, p. 80 ; R. J. Avanesov *(op. cit.)*, pp. 132-133.

[33] A. Martinet, *Phonology as Functional Phonetics*, London, 1949, pp. 9-10 ; R. Jakobson, C. G. M. Fant, M. Halle, *Preliminaries to Speech Analysis = Technical Report No. 13* (1952), MIT Acoustics Laboratory; J. Lotz, « The Structure of Human Speech », *Transactions of the New York Academy of Sciences*, ser. II, 16.373-384 (1954), p. 383.

[34] R. Jakobson and M. Halle, *Fundamentals of Language*, The Hague 1956, p. 9.

[35] "Die Kombinatorischen Varianten sind also eben deshalb möglich, weil verschiedene phonetische Eigenschaften desselben Lautes in verschiedenen Funktionen verwendet werden," "Über eine neue Kritik des Phonembegriffes," *Archiv für vergleichende Phonetik* I, 3, p. 137 (1937).

convergence which were established in the course of a common historical development and contact between dialects. Let us consider, for example, the configurational feature of stress in the areas where stress is not distinctive. All Polish dialects (except the linguistic islet of Podhale) have in common the stress on the penultimate syllable, whereas most Czech, Moravian and Slovak dialects have an initial stress. Comparing, furthermore, forms which are phonetically identical in various dialects, such as Polish [drog′i] 'dear', [mǫš] 'husband', we observe that in some dialects [g′], [š], [ǫ] are contextual variants of the phonemes /g/, /ž/, /o/, whereas in others /g′/, /š/, /ǫ/ are in phonemic contrast with /g/, /ž/, /o/. Examples of such relations are easily found when we reinterpret phonemically the isoglosses of existing dialect-maps. An interesting example is supplied by the so-called "Mazur" dialects in Northeastern Poland (in Areas III and IV on map 2). In the Western "Mazur" area, the velar consonants have redundant palatalization before front vowels including /æ/, but in the Eastern area, where /æ/ coalesced with /a/, the palatalization of velars is distinctive; we find here such interesting pairs as /droga/ 'dear' and /drog′a/ 'road', /suxa/ 'dry' and /mux′a/ 'fly'.[36] The variant [æ] which is found after palatals in some Eastern Slovak areas, brings these dialects closer to the Central Slovak dialects in which /æ/ is phonemic.[37] Instead of obscuring the homogeneity of phonetic substance, a structural approach can emphasize its different functional utilization across various dialect areas. Over maps showing discrete phonemic systems, we could plot maps of phonically uniform territories.

8. The above remarks concerning the study of dialects from the point of view of discreteness as well as of continuity, have further implications if we pass from the synchronic aspect of dialect studies to its diachronic aspect. Dialectology, which has been one of our main sources of information on historical processes, reveals intimate connections between synchrony and diachrony. As in the case of stylistic alternants which pertain both to the "vertical" time axis and to the "horizontal" space axis, the different utilization of features across linguistic areas reflects the phenomena which have been known in diachronic linguistics

[36] K. Nitsch, *Wybór pism polonistycznych*, III, *Pisma pomorzoznawcze*, Wrocław, 1954, p. 346.

[37] Z. Stieber, "Stanowisko mowy Słowaków" *(op. cit.)*, pp. 49 ff.

as "dephonemicization," "phonemicization" and "rephonemicization."[38]

The allophonic status of nasal vowels in a number of Polish dialects reflects the tendency in the history of Polish (and more remotely in other Slavic languages) to eliminate nasal vowels. The phonemic status of the palatalized consonants /k′, g′/ is a structural innovation which took place primarily in the Western Polish dialects. Until recently /k′, g′/ were in a state of oscillation between a phonemic and allophonic status in the literary language also. The phonemic value of /γ/ in the Central Russian dialects, in which it is opposed to /x/ as voiced vs. voiceless continuant, can be cited as an example of an areal correspondence to the historical process of rephonemicization. In other Russian dialects where [γ] has preserved its original status as a variant of /g/, the phoneme /x/ is opposed to both /g/ and /k/ as a continuant vs. stops and its voiceless quality is redundant.

The correlation between zonal distribution and temporal succession has its limitations in the case of dialectal contact, where speakers are known to take over distinctive as well as non-distinctive features from their neighbors. But as diachronic linguistics, with its emphasis on immanent change, takes into account the impact of external factors on the history of a language, dialectology cannot ignore problems of dialect interference and, instead of seeking in space a mirror-image of time, it should rather try to indicate the problems which dialectology has in common with diachronic linguistics. As the formation of culture languages or of other vehicles of intercommunication is often achieved through the blending of different systems, the "mixed" dialects may overlap with the "core" dialects at some points or on some levels of their phonemic systems, but they lack the overall complex of features which enables us to view the "core" dialects as varieties of certain phonemic types.

If we press further the analogy between the two branches of comparative linguistics, the problem of grouping dialects into an overall system has its correlate in the historical linguist's quest for periodization, i.e. for establishing a turning point (or points) in the history of a language where the structural transformation ("Umbruch") of a system was far more profound and significant than changes in its inventory which left the phonemic correlations intact. It is at such junctures of historical development that

[38] R. Jakobson, "Prinzipien der historischen Phonologie," *TCLP* 4.241 ff. (1931).

certain dialects acquire or lose distinctive features or combinations of features which set them off from other related dialects.

As diachronic linguistics was able to avoid the dead end of early structuralism by abandoning its concept of isolated systems juxtaposed statically in time, dialectology can resist the danger of atomization by studying dialectal varieties in their two-fold aspect of discreteness and continuity.

Indiana University.

NOTE ON THE DEVELOPMENT OF THE MODERN FRENCH VOWEL SYSTEM

ALEXANDER HULL

Ferdinand Brunot, discussing the phonetics of eighteenth-century French, makes the following remark:

> Au XVIII^e siècle le renversement qui avait commencé au XVI^e siècle est complet. Il n'y a plus à la base une langue parlée, que l'écriture reproduit, en essayant d'en suivre de près ou de loin les variations... Ce qui est à la base, c'est une langue écrite, fixée ou qu'on croit fixée par les textes des écrivains, les observations des grammairiens et des lexicologues, et que reproduit bien ou mal la parole.
>
> On peut déterminer d'après cela quel sera le mouvement phonétique. Les causes naturelles qui le produisent ne sont pas détruites ; elles ne peuvent pas l'être, étant des forces de la nature. Mais elles sont contraintes et souvent contrariées.
>
> Les sons se modifient ; ils ne se transforment plus. Pour parler avec plus de précision scientifique, ils s'altèrent encore dans la mesure où ils peuvent le faire sans sortir d'un même timbre, pas au delà...[1]

This statement, which suggests that the whole phonemic structure of the modern French spoken language has been molded by its orthography, has significant implications for the study of the history of all languages possessing an important literature and codified by grammarians. The purpose of this article is to examine the background of Brunot's assertion in the light of contemporary theories of diachronic phonemics, and with reference to certain developments which have taken place in the vowel system of Canadian French, in a dialect which was almost completely cut off from the written language for 150 years.

[1] *Histoire de la langue française*, VI (Paris, 1952), 973.

Brunot's thesis is that the vowel E was tending toward A, but was prevented from changing by the influence of the spelling, except in the case of those words spelled in *oi*, where neither pronunciation would correspond to the written form.[2] Let us then examine the vowel system of French in 1700, to determine whether the tendency to change E to A has any structural basis. André Martinet, examining the grammatical treatise of Gile Vaudelin,[3] arrives at the following pattern of oral vowels (long and short) for the language described by this author (p. 17):

i	ũ	u
ę	ö	o
ę̨		
	a	

It is obvious that this is an asymmetrical system—one, then, which contains the seeds of possible readjustment. It is also clear that the simplest ways to restore symmetry would be either by the confusion of /ę/ and /ę̨/, or /ę/ and /a/, or by a movement of /a/ to a back position, which would permit the opening of /ę/ to /a/ without a phonemic merger. The confusion of the two E phonemes in closed syllables seems already to have taken place at this time.[4] It would appear also that some provincial dialects did not make the distinction between the two E's in open syllables.[5] The standard language did not lose this distinction, however, until the recent tendency to confusion started, presumably as the result of similar structural pressures. Nor did the other possible changes take place, although the change of /wę/ to /wa/ and the confusion of /ę/ and /a/ in certain cases before /r/ have been suggested as indicating a very open pronunciation of /ę/, which may have been tending to merge with /a/ as intimated by Brunot.

Let us now, however, examine the vowel system of the French-Canadian dialect spoken in Windsor, Ontario and the surrounding region. This dialect, which was formed between 1700 and 1750, largely by the migration of second-generation Canadians from the Montreal area, has been relatively unaffected by changes filtering

[2] *Ibid.*, pp. 973-974.

[3] André Martinet, « Notes sur la phonologie du français vers 1700 », *Bulletin de la Société Linguistique de Paris*, XLIII (1946), 13-23.

[4] Cf. Th. Rosset, *Les Origines de la prononciation moderne* (Paris, 1911), pp. 117 ff.

[5] See Ch. Thurot, *De la prononciation française depuis le commencement du XVI^e siècle* (Paris, 1881-1883), I, 50 ff.

into Quebec from the standard language which may distort the original picture there.[6] It will be necessary to present separately the system of stressed vowels in open and in closed syllables. The system of oral vowels in open syllables is as follows (there are also nasal vowels /ã/, /ẽ/, and /õ/) :

i	ü	u
e	ö	o
a		å

Here the phoneme /a/, the pronunciation of which ranges from a fairly front sound, almost [æ], to a low central [a], represents an original final /ę/, as in *j'étais* /žeta/, *vrai* /vra/, *forêt* /fora/. The phoneme /å/, low back rounded, represents original /a/, as in *ral* /rå/, *bas* /bå/, *Canada* /kanadå/.

In closed syllables, the system is as follows:

i		ü	u	
e	e:	ö	o	o:
	a		å	

The phoneme /å/ could be interpreted here as representing /a:/, but in view of the evident distinction of the two phonemes in open syllables, this does not seem justified. The long vowels /e:/ and /o:/, as well as /å/, tend to diphthongize. Here the vowel /a/ represents original /a/, original /ę/ being preserved as /e/,[7] due no doubt to the fact that /ẹ/ and /ę/ were already merged in this position before the emigration of the Canadians. For example, *fait* (past participle) /fet/, pronounced [fęt], where the consonant has been preserved, is distinguished from *il fait* /fa/. Compare *fête* /fe:t/, pronounced [fẹ:t], [feįt], [faįt]; *patte* /pat/; *pâte* /påt/, pronounced [på:t] or [påųt].

At this point it is of interest to note the development in this dialect of the words with original /wę/. In open syllables, three different results occur: /we/, /wa/, and /wå/.[8] Of these, the second is the most general. Only a few words (*moi*, *toi*, *je vois*, *je bois*) have /we/, and this alternates with /wa/. It probably originated in a provincial pronunciation.[9] The words with /wå/

[6] Cf. the author's article in *Orbis*, V (1956), 35-60.

[7] Except before /r/ plus consonant, where /ę/ has become /a/.

[8] I omit those words which show /e/ for /wę/, as *froid* /fret/, *je crois* /kre/, obviously of provincial origin.

[9] Cf. Thurot, I, 355.

are a group spelled with final *-s* and *-x* (*bois*, *pois*, *noix*, etc.) which undoubtedly had the /a/ sound long before the other words in France.[10] The normal /wa/ development, in such words as *doigt*, *roi*, *oie*, results in a pronunciation identical to the standard in these words. In the Windsor dialect, however, this sound is phonemically identical to the one in *forêt*, not the one in *rat*. It must have been a separate development in Canada from an original /wę/, as we can judge by the retention of /wę/ in certain other dialects of North American French where final /ę/ has not become /a/ (e.g., in Ste. Genevieve, Missouri, a place even more isolated than Windsor[11]). The change of /wę/ to /wa/ has come about, then, not in isolation, but as part of the wider change from /ę/ to /a/. This is also proved by the fact that it did not take place in closed syllables (*avoine* /a(v)wen/, *boîte* /bwe:t/, etc.).

In this Canadian dialect, then, the asymmetry of the seventeenth-century French vowel system has been resolved by the shift of /a/ to /å/ and /ę/ to /a/. The change of /wę/ to /wa/ appears as a part of this general resolution, which is limited to open syllables. One may infer that this might have been the natural tendency of the language in France, too, as the most economical development which would avoid phonemic mergers, and that it was frustrated by the reaction of educated people influenced by the writing system, resulting in the separation of the *oi* words from the remainder of the words in /ę/. But it would be dangerous to assert this too positively. Too many other factors may have intervened. The question of the functional yield of the opposition ẹ/ę must be considered. It is to be noted that not even all North American French dialects made the change. The two E's have merged in the Acadian of Louisiana, which has also shortened the long vowels without creating any additional distinctions.[12] Other explanations are possible for the change of /wę/ to /wa/ in France, which may have no connection with the Canadian change. The /ę/ in the combination /wę/ was probably more open than elsewhere. In this connection, it is interesting that the /wa/ of Windsor is pronounced with a vowel farther back than normal for the /a/ phoneme, although phonetically identical to the /a/ in /ar/ and /ra/, and definitely distinct from /å/.

[10] *Ibid.*, I, 356 ff.

[11] See J.-M. Carrière, "The Phonology of Missouri French: A Historical Study," *French Review*, XIV (1941), 410-415 and 510-515.

[12] Cf. Wm. A. Read, *Louisiana-French* (Baton Rouge, 1931), p. xx.

What we can assert is this: the orthography, and the reaction of certain influential speakers to it, may have played a part in determining the direction of development of the French vowel system, favoring the eventual restoration of symmetry by means which were not the most economical ones at the time. The tradition of the pronunciation of the letter *e* with an E sound could not be lost as long as no change was taking place in closed syllables. It would be interesting to compare the situation in English, where the great vowel-shift, taking place at an earlier date, was not prevented by orthographical considerations. In English today, with the multiplicity of sounds represented by each vowel letter, there is no tradition to prevent further changes. In French, on the other hand, this tradition remains powerful, and it is not likely that the standard language will easily accept changes which would destroy the relationships between the letters and the sounds as they now exist. Such factors must at least be considered when discussing the history of the phonemic structure of a literary language.

University of Massachusetts.

THE PHONOLOGY OF THE WORD IN MODERN STANDARD MONGOLIAN

Don Graham Stuart
with the collaboration of
Matthew M. Haltod

This paper describes the phonology of the word in Modern Standard Mongolian as realized in the usage of Matthew Magatburin Haltod, our co-worker in the research leading up to the present report. What is intended here by the term Modern Standard Mongolian is the supra-dialectal form of language, based upon the Xalxa dialect, which has developed with Mongolian nationalism as the vehicle of cultural life in present day Mongolia. Although this usage is not of recent development, there having existed a semi-colloquial literature in the Mongolian script for some time, its unification—especially in phonological matters—has certainly been accelerated by the adoption of a Cyrillic orthography that adequately reflects the pronunciation. Its widespread use by non-Xalxa speakers precludes our identifying it completely with the Xalxa dialect; the usage here described is that of a Čaxar speaker whose childhood language was the Čaxar dialect and who has been educated in Manchu, Chinese, Tibetan, and Japanese. Some deviations from Xalxa usage are therefore to be expected.

M. M. H. was born in 1917 in the Čaxar Aimag[2] at Šuluun xöx Xušuu, about 700 *li* from Kalgan. He was taught in Mongolian,

[1] This paper was prepared in connection with the project work of the Linguistic Survey of Asia and was made possible by funds granted by the Ford Foundation. That Foundation is not, however, the author, owner, publisher, or proprietor of this paper and is not to be understood as approving by virtue of its grant any of the statements made or expressed therein.

[2] Cyrillic forms are cited herein in a transcription which differs somewhat from the system of *The American Slavic and East European Review* but which corresponds to widespread practice and is, therefore, unlikely to cause confusion. The Cyrillic

Manchu, and Tibetan by his father and by tutors. Later, he learned Mandarin from his elder brother. From April 1937 to July 1939 he studied in Japan and speaks Japanese with complete fluency. Japanese was used as the working language in the collaboration leading up to the present report.

N. Poppe[3] distinguishes three subgroups in East Mongolian: 1) Burjat, 2) Xalxa, and 3) a southern complex consisting of a) Čaxar, b) Urat, c) Xarčin, d) Ordos, e) Xorčin, and f) Bargu. M. M. H. states that of these Burjat—unless 'spoken to be understood by outsiders'—is completely unintelligible to Čaxar speakers, whereas Xalxa and Čaxar are mutually understandable without difficulty. Čaxar speakers, unless striving for elegance, will not bother to modify their manner of speaking in addressing a Xalxa speaker. Although aware that the standard language is being formed on the Xalxa dialect of Ulaan Baatar (/Ulaŋ Baatr/), Čaxar speakers are able to distinguish broad dialect variations within Xalxa and are reluctant to admit that the standard is identical with any regional usage.

PRELIMINARY REMARKS

We deal here exclusively with the phonology of the word. On the phonological level, the word in Mongolian is characterized by a) its intrinsic phonemic shape and b) by a culminative accent. These two criteria fail to coincide in some cases, e.g. there are some few cases where two complexes which by their phonemic shape cannot belong to the same word are included in the envelope of a single culminative accent. The word as such is further delimited by various boundary phenomena; however, such of

alphabet in this transcription is as follows, the transcriptions of the special letters employed for Mongolian being given in parenthesis: *a, b, v, g, d, je, jo, ž, z, i, j, k, l, m, n, (o), ö, p, r, s, t, u, (ü), f, x, c, č, š, šč, z, y, b, e, ju, ja.* In the Cyrillic, the *ö* is written with an *o* with a horizontal line in it, like theta, and the *ü* with a *y* which is differentiated from the sign for *u* by being more angular and having a straight, rather than looped, tail. We were tempted to effect a typographical economy by using *y* for the weak jer, and *w* for the sign of nonpalatalization, but in the end diffidence prevailed against our making this radical, but sensible, break with tradition. Unless otherwise indicated, all Mongolian names etc. are given in transliteration of the modern orthography. Following custom, transliterated forms are given herein in italics, phonemic transcriptions between diagonal bars, and phonetic approximations between square brackets.

[3] *Khalka-Mongolische Grammatik.* Wiesbaden, 1951. (Akademie der Wissenschaften und der Literatur: Veröffentlichungen der orientalischen Kommission, Band I.) p. 3.

these as are not associated with the rise and fall of the culminative accent are largely facultative and are therefore highly unreliable as defining criteria.

The phonemic shape of words in Mongolian is being radically influenced by foreign words which are largely accepted without much modification as long as an approximation in Mongolian phonemes is possible, even though an unusual phoneme sequence is produced. The so-called vowel harmony is, of course, neglected more often than not in recent loan words, at least among educated persons. The extent to which foreign loans are assimilated to the Mongolian vowel harmony in vulgar usage is a question which must be investigated. In the present study, only indigenous words or thoroughly assimilated loans are considered; this may, however, yield a phonological system which will be inadequate for the analysis of all texts. Some thoughts on this question will be presented at the conclusion of the report.

The theoretical approach adopted here involves no departure from that of the later "Prague School." It will be obvious that we have closely followed the model established by André Martinet in his treatise *La description phonologique.*[4]

It is a difficult question whether vowel harmony should be treated under the rubric of phonology or morphophonology. We have herein set forth the general rules of the vowel harmony in the section on the distribution of phonemes. It is impossible however to accept the view that vowel harmony is the result of some kind of inevitable prohibition on the alternation of members of certain oppositions of distinctive features within a word arising from the physical or acoustical nature of the sound features. We must also reject the related viewpoint[5] that certain vowel features may best be regarded as suprasegmental. It is decidedly wrong to think of vowel harmony as a kind of euphony; Mongol speakers have no difficulty in pronouncing words in which vowel harmony is contradicted. Exceptions occur not only in foreign loan words but also with native roots in some cases where the same culminative accent incorporates more than one minimal sign, e. g. *javžee* /jawǰä/ 'gone'; *javaxgüj* /jawxgüi/ 'not going'. On the other hand, the phonetic value of the 'reduced vowels', the indefinite vocalizations that relieve difficult consonant clusters, does seem to be determined by the vocalic environment.

[4] *La description phonologique, avec application au parler franco-provençal d'Hauteville (Savoie)*, 1956.

[5] The view, apparently, of E. P. Hamp, cf. S. Hattori in article cited in f. n. 7

LITERATURE

The literature devoted to the phonology of Modern Mongolian is not very extensive. The excellent grammar of N. Poppe (*Khalka-Mongolische Grammatik*, see f. n. 3) gives 19 pages to "Lautlehre" but cites no minimal pairs and omits discussion of questions of interest to the phonologist. Unfortunately, it is impossible to compare Prof. Poppe's conclusions with those reached here because the usages described are not identical. The same may be said of the 31 page *Fonetika* in B. X. Todajeva's *Grammatika sovrjemjennogo Mongol'skogo jazyka.*[6]

Studies in manuscript by William M. Austin ('Spoken Mongol') and Edvard P. Torjesen ('Mongol Phonemic Romanization') are known to us through the discussion of their conclusions in an article by S. Hattori in *Gengo Kenkyû.*[7] Hattori's article is full of interesting suggestions and stimulating insights; however, since it is chiefly a critique of the above cited papers, his own views are not set forth as fully as it may be hoped they will be sometime in the future.

A recent text-book by William M. Austin, Gombojob Hangin, and Urgunge Onon, entitled *A Mongol Reader* (American Council of Learned Societies, Washington, D. C., 1956, pp. [ii] 248), gives some indications under 'Mongol Writing and Mongol Sounds' and 'Vowel Harmony', but despite the use of the terms 'phonemes' and 'allophones' does not go far beyond the sort of explanation of pronunciation that one expects to find in pocket phrase books.

Useful information on the Cyrillic orthography is to be found in Nomura Tadayoshi's contribution on Mongol in *An Introduction to the Languages of the World*, Vol. II, p. 555 (Kenkyûsha, Tôkyô, 1955, in Japanese). He gives credit to a book by Sakamoto Koretada: *Mongol ügijg surax bičig*, Tôkyô Gaikokugo Daigaku, 1951, and an article by Aomatsu Genichi: 'Shinmoji ni yoru Gaimôkogobunten no tokuchô to jûrai no bunten to no hikaku' *Ôsaka Gaikokugo Daigaku Gakuhô* I, 1952. Neither of these has been available to us.

[6] Moscow, 1951.

[7] 'Phonemic Structure of Mongol (Chakhar Dialect)'. *Gengo Kenkyû, Journal of the Linguistic Society of Japan*, Nos. 19-20, 1951, pp. 68-102. (In Japanese).

PROSODY OF THE WORD

The accent in Modern Mongolian is culminative, without distinctive function, its position being determined as follows:—In monosyllables, the accent falls on the syllabic vowel; in polysyllabic words, the accent falls on the first long vowel or diphthong, or if these be lacking, on the second syllabic vowel other than short /i/; in disyllabic words with short /i/ in the second syllable, the first syllabic vowel takes the accent (all vowels are syllabic except /i/ immediately before or after another vowel and /u/ in the rare combinations /ua/ and /uai/ after an initial consonant in Sinic loan words).

THE PHONEMES

I. *The Consonants*

1. The phoneme /b/.

This phoneme commutes with other consonantal phonemes in word initial position and non-initially after another consonant, but not post-vocalically. Its phonological identity results from the oppositions demonstrated in the following minimal pairs

1: (b/m) ***bal*** /bal/ 'honey'∽***mal*** /mal/ 'cattle, domestic animal'

2: (b/w) ***bandan*** /baŋdŋ/ 'bench, seat'∽***vandan*** /waŋdŋ/ 'a kind of silk offering cloth.'

The phonemes /b/ and /w/ are diachronically the result of a phonemic splitting; /w/ corresponds to the weak post-vocalic articulation of the old phoneme. The phonematization of this opposition occurred initially to accommodate loan words and non-initially as a result of the phonological loss of reduced vowels which brought /b/ and /w/ in such functions as *-CbV* and *-CVw* (>*-Cb* and *-Cw*) into opposition in post-consonantal position.

This opposition seems to be neutralized after /m/, i.e. such words as ***dombo*** /domb/ 'tea can, tea pot'; ***sambar*** /sambr/ 'wooden book covers (for Mongolian style binding)' occur, but */*-mw-*/ is not to be found.

3: (b/d) ***ber*** /bär/ 'bride'∽***der*** /där/ 'pillow'.

4: (b/g) ***bar*** /bar/ 'tiger'∽***gar*** /gar/ 'hand'.

2. The phoneme /m/.

The phonological identity of this phoneme results from the following comparisons:

1: (m/b) cf. 1. 1.
2: (m/w) *maar* /maar/ 'salt marsh'∾*vaar* /waar/ 'tile' *sam* /sam/ 'comb'∾*sav* /saw/ 'container (for food, etc.)'.
3: (m/n) *mad* /mad/ (a special word in chess terminology)∾ *nad* /nad/ 'to me'; *lam* /lam/ 'a Buddhist priest'∾*xana* /xan/ 'wall'.
4: (m/ŋ) [opposed only in syllabic or word final position] *zam* /zam/ 'road'∾*zan* /zaŋ/ 'nature, disposition, habit'.

3. The phoneme /w/.

The phonological identity of this phoneme results from the following comparisons.

1: (w/b) cf. 1.2.
2: (w/m) cf. 2.2.

4. The phoneme /t/.

The phonological identity of this phoneme results from the following comparisons:

1: (/t/d) *tarax* /tarx/ 'deploy, be dispersed' ∾*darax* /darx/ 'to print, press'; *at*/ at/ 'a castrated camel' ∾*ad*/ad/ 'a demon'.
2: (t/n) *xat* /xat/ 'be firm'. (imp.)' ∾*xana* /xan/ 'wall'; *xatax* /xatx/ 'to become firm, hard' ∾*xanax* /xanx/ 'to satisfy'; *ter* /tär/ 'that' ∾*ner* /när/ 'name'.
3: (t/s) *bat* /bat/ 'hard' ∾*bas* /bas/ 'again'; *tav* /taw/ 'five' ∾*sav* /saw/ 'container'.
4: (t/č) *tamaas* /tamas/ 'from Hell' ∾*čamaas* /čamas/ 'from you'; *xojt* /xoit/ 'northern' ∾*xojč* /xoič/ 'later'.
5: (t/c) *tas* /tas/ 'vulture' ∾*cas* /cas/ 'snow'; *at* /at/ 'camel' ∾ *ac* /ac/ 'large rake'.
6: (t/l, r) *tav* /taw/ 'five' ∾*lav* /law/ 'surely'; *bat* /bat/ 'solid' ∾ *bal* /bal/ 'honey' ∾*bar* /bar/ 'tiger'; *at* /at/ 'camel ∾*ar* /ar/ 'posterior' (adj.).

5. The phomene /d/

The phonological identity of this phoneme results from the following comparisons:

1: (d/t) cf. 4.1.

2: (d/b) cf. 1.3.

3: (d/n) *xad* /xad/ 'a rocky ridge' ∽*xana* /xan 'wall'; *dam* /dam/ 'intermediary, broker' ∽*nam* /nam/ 'short and broad'.

4: (d/g) *dal* /dal/ 'seventy' ∽*gal* /gal 'fire'; *ted* /täd/ 'they' ∽*teg* /täg/ 'do thus! (imp.)'.

5: (d/z) *del* /däl/ 'mane (of a horse)' ∽*zel* /zäl/ 'halter'; *edgüj* /ädgüi/ 'propertyless' ∽*ezgüj* /äzgüi/ 'absent from home'.

6: (d/l, r) *dam* /dam/ 'intermediary' ∽*lam* /lam/ 'a Buddhist priest'; *dax* /dax/ 'mantel, overcoat' ∽*rax* /rax/ [name of an evil deity]; *ad* /ad/ 'a demon' ∽*al* /al/ 'kill! (imp.)' ∽*ar* /ar/ 'posterior' (adj.).

6. The phoneme /n/.

The phonological identity of this phoneme results from the following comparisons:

1: (n/t) cf. 4.2.

2: (n/d) cf. 5.3.

3: (n/m) cf. 2.3.

4: (n/ŋ) [neutralized in syllabic initial position] *xaana* /xaan/ 'in bed' ∽*xaan* /xaaŋ/ 'emperor' (<Old Mongol *xaγan*).[8]

5: (n/l) *nam* /nam/ 'political party, faction' ∽*lam* /lam/ 'a Buddhist priest'; *ene* /än/ 'this' ∽*el* /äl/ 'all'.

7. The phoneme /s/.

The phonological identity of this phoneme results from the following comparisons:

1: (s/t) cf. 4.3.

2: (s/š) *es* /äs/ 'cancellation' ∽*eš* /äš/ 'tree trunk'; *sar* /sar/ 'moon, month' ∽*šar* /ʃar/ 'yellow'.

Before /i/ or /ii/ the opposition of /s/ ∽ /š/ is very seldom realized. M. M. H. could not find any examples of this opposition before short /i/ except the names of the letters /šii/ and /sii/ in the Mongol script pronounced in the traditional manner of reciting the alphabet. Before long /ii/ only the following near-pair could be found; it,

[8] Cf. Karl H. Menges : *The Oriental Elements in the Vocabulary of the Oldest Russian Epos, the Igor Tale*, Supp. to *Word*, Vol. 7 (1951), pp. 32-35, for a discussion of the etymology of this word

however, seems decisive: ***sijlex*** /siilx/ 'carve in relief, engrave' ***šijdex*** /šiidx/ 'to decide'.

3: (s/c) ***sam*** /sam/ 'comb' ∽***cam*** /cam/ 'a Buddhist dance'; ***as*** /as/ 'straight' ∽***ac*** /ac/ 'large rake'.

4: (s/x) ***sar*** /sar/ 'moon' ∽***xar*** /xar/ 'black'; ***as*** /as/ 'straight' ∽ ***ax*** /ax/ 'elder brother'.

8. The phoneme /š/.

The phonological identity of this phoneme results from the following comparisons:

1: (š/s) cf. 7.2.

2: (š/x) ***šar*** /šar/ 'yellow/ ∽***xar*** /xar/ 'black'; ***eš*** /äš/ 'handle' ∽ ***ex*** /äx/ 'mother'.

3: (š/č) ***šil*** /šil/ 'crow' ∽ ***čil*** /čil/ (in compound ***čil nogoon*** /čil nogoŋ/ 'deep green'); ***biš*** /biš/ (neg. particle) ∽***bič*** /bič/ 'monkey'.

4: (š/r) ***eš*** /äš/ 'handle' ∽***er*** /är/ 'male'.

9. The phoneme /c/.

The phonological identity of this phoneme results from the following comparisons:

1: (c/t) cf. 4.5.

2: (c/s) cf. 7.3.

3: (c/z) ***car*** /car/ 'hard crust on the surface of snow' ∽ ***zar*** /zar/ 'order, proclamation, public announcement'; ***ac*** /ac/ 'large rake' ∽***az*** /az/ 'luck, fortune, fate'.

4: (c/č) ***cadax*** /cadx/ 'satisfy' ∽***čadax*** /čadx/ 'can, be able'; ***ac*** /ac/ 'large rake' ∽***ač*** /ač/ 'grandson'.

10. The phoneme /z/.

The phonological identity of this phoneme results from the following comparisons:

1: (z/d) cf. 5.5.

2: (z/c) cf. 9.3.

3: (z/ǰ) ***zar*** /zar/ 'proclamation' ∽ ***žar*** /ǰar/ 'sixty'; ***az*** /az/ 'fortune, luck' ∽***až*** /aǰ/ 'work, task'.

11. The phoneme /č/.

The phonological identity of this phoneme results from the following comparisons:

1: (č/t) cf. 4.4.
2: (č/š) cf. 8.3.
3: (č/c) cf. 9.4.
4: (č/ǰ) *čil* /čil/ (in *čil nogoon* /čil nogoŋ/ 'deep green') ∾ *žil* /ǰil/ 'year'; *ač* /ač/ 'grandchild' ∾*až* /aǰ/ 'task, work'.

12. The phoneme /ǰ/.

The phonological identity of this phoneme results from the following comparisons:

1: (ǰ/z) cf. 10.3.
2: (ǰ/č) cf. 11.4.
3: (ǰ/d) *žuuč* /ǰuuč/ 'intermediary, broker' ∾*duuč* /duuč/ 'singer'; *až* /aǰ/ 'task' ∾*ad* /ad/ 'demon'.

13. The phoneme /g/.

The phonological identity of this phoneme results from the following comparisons:

1: (g/b) cf. 1.4.
2: (g/d) cf. 5.4.
3: (g/ŋ) *tug* /tug/ 'banner ∾*tuŋ* /tuŋ/ 'really'.
4: (g/x) *gar* /gar/ 'hand' ∾*xar* /xar/ 'black'; *zag* /zag/ (name of a kind of plant) ∾*zax* /zax/ 'boundary'.

14. The phoneme /ŋ/.

The phonological identity of this phoneme results from the following comparisons:

1: (ŋ/m) cf. 2.4.
2: (ŋ/n) cf. 6.4.
3: (ŋ/g) cf. 13.3.
4: (ŋ/l) *zun* /zuŋ/ 'summer' ∾*zul* /zul/ 'a light, lamp'.
5: (ŋ/x) *an* /aŋ/ 'hunting, the chase' ∾*ax* /ax/ 'elder brother'.

15. The phoneme /x/.

The phonological identity of this phoneme results from the following comparisons:

1: (x/š) cf. 8.2.
2: (x/g) cf. 13.4.
3: (x/ŋ) cf. 14.5.
4: (x/s) *xar* /xar/ 'black' ∾*sar* /sar/ 'moon, month'; *ex* /äx/ 'mother' ∾*es* /äs/ (neg. part.) 'not'.

16. The phoneme /l/.

The phonological identity of this proneme results from the following comparisons:

1: (l/t) cf. 4.6.
2: (l/d) cf. 5.6.
3: (l/n) cf. 6.5.
4: (l/ŋ) cf. 14.4.
5: (l/r) *bal* /bal/ 'honey' ∾*bar* /bar/ 'tiger'; *xolʲ* /xoli/ 'mix! (imp.)' ∾*xorʲ* /xori/ 'twenty'.

17. The phoneme /r/.

The phonological identity of this phoneme results from the following comparisons:

1: (r/t) cf. 4.6.
2: (r/d) cf. 5.6.
3: (r/š) cf. 8.4.
4: (r/l) cf. 16.5.

II. *The Archiphoneme /I/.*

18. The archiphoneme /I/ results from the neutralization of the opposition between the phonemes /i/ and /j/. The following positions need be distinguished:

a. In word initial position the opposition /i/ ∾/j/ is neutralized, and only the archiphoneme is realized. We shall adopt the convention of writing *j*. In the Cyrillic orthography, the Russian compound vowels are used, e.g. *jaarax* /jaarx/.

b. As the second element of diphthongs, only the archiphoneme is realized. We shall represent it herein with an *i*, e.g. *bajna* /bain/.

c. As the central vowel of a syllable, only the archiphoneme is realized both in initial and following syllables. Long and short quantity is distinguished in all syllabic positions. We shall adopt the convention of writing /i/ and /ii/, e.g. *bi* /bi/, *širex* /širx/, *morʲ* /mori/, *amʲt* /amit/, *salxin* /salxiŋ/.

d. In postconsonantal prevocalic position, /i/ and /j/ are opposed, e.g. *xiarax* /xiarx/ ∾ *xjarax* /xjarx/; *baria* /baria/ ∾ *barʒja* /barja/ (cf. also *barʲja* /barija/).

In initial position the /I/ archiphoneme is opposed to other consonantal phonemes. The following comparisons contribute to its phonological identity:

1: (I/b, m, t, d, g) *jal* /jal/ 'guilt' ∾*bal* /bal/ 'honey' ∾ *mal* /mal/ 'cattle' ∾*tal* /tal/ 'steppe' ∾*dal* /dal/ 'seventy' ∾ *gal* /gal/ 'fire'.

In other positions its phonological identity results from comparisons such as the following:

2: (I/a) *im* /im/ 'mark of ownership on domestic animal's ear' ∾*am* /am/ 'mouth'; *širex* /širx/ 'decorative stitching or embroidery on felt' ∾*šarax* /šarx/ 'to roast (meat), to be burned by the sun, to sun bathe'; *barix* /barix/ 'to hold' ∾ *barax* /barx/ 'to finish up, use up (wastefully?)'.

3: (I/o) *id* /id/ 'bodily strength' ∾*od* /od/ 'star'; *činij* /činii/ 'your (poss. pn.)' ∾*čonyn* /čoniiŋ/ (čono- /čon-/) 'of the wolf'.

4: (I/u) *id* /id/ 'bodily strength' ∾*ud* /ud/ 'a kind of weeping willow tree'; *bid* /bid/ 'we' ∾*zud* /zud/ 'cattle famine in winter (due to grazing places being snowed-over)'.

5: (I/a) *ir* /ir/ 'blade' ∾*er* /är/ 'male, masculine'; *xil* /xil/ 'frontier' ∾*xel* /xäl/ 'tongue; language'.

6: (I/o) *il* /il/ 'clear' ∾*öl* /öl/ 'nutrition, nourishment'; *šine* /šin/ 'new'∾ *šönö* /šön/ 'night'.

7: (I/ü) *il* /il/ 'clear, not concealed' ∾*ül* /ül/ 'not' (verbal prefix); *šid* /šid/ 'sew!' (imp.) ∾*šüd* /šüd/ 'teeth'; *xedijl* /xädiil/ 'how much, how many' ∾ *xedüül* /xädül/ 'how many persons'.

The opposition of /i/ and /j/ in postconsonantal prevocalic position results from such comparisons as follow:

8: (i/j) *xiarax* /xiarx/ 'to wave in the wind (of grass)' ∾ *xjarax* /xjarx/ 'to comb' (*nomen futuri* /xjar-/) [translation not certain]; *baria* /baria/ 'a halter for a cow (usually attaching to the horns)' ∾ *barʒja* /barja/ 'look, here it is; *voilà*'. (NB /j/ has no palatalizing effect on preceding consonant and vowel; *barʲja* /barija/ 'let us seize (it)' should be added to the comparison).

The opposition of length realized in the occurrence of /I/ as a central vowel results from the following comparisons:

9: (I/II) *im* /im/ 'mark of ownership on a domestic animal's ear' ∾*ijm* /iim/ 'in this manner'; *šidex* /šidx/ 'to sew on,

baste, stick up' ∾*šijdex* /šiidx/ 'to decide definitely, to determine'; *či* /či/ 'thou' ∾*čij* /čii/ 'a Chinese foot (unit of measure)'.

III. *The Vowels.*

19. The phoneme /a/.

The phonological identity of this phoneme results from the following comparisons:

1: (a/I) cf. 18.2.

2: (a/o) *al* /al/ 'kill!' (imp.) ∾*ol* /ol/ 'find (it); look (for it)! (imp.)'; *bal* /bal/ 'honey' ∾*bol* /bol/ (particle designating subject) ; *xorь* /xori/ 'twenty' ∾ *xarь* /xari/ 'foreign'.

3: (a/u) *at* /at/ 'a castrated camel' ∾*ut* /ut/ 'smoke out! (imp.; as devils with incense)'; *bat* /bat/ 'constant, firm, loyal' ∾ *but* /but/ 'bush'.

4: (a/ä) *ax* /ax/ 'elder brother' ∾*ex* /äx/ 'mother'; *xad* /xad/ 'rock, cliff' ∾*xed* /xäd/ 'how many'; *xana* /xan/ *term. lec.* for the latticed wall of a Mongol yurt ∾ *ene* /än/ 'this'; *tad* /tad/ 'you all' ∾*ted* /täd/ 'they all'.

5: (a/ö) *ar* /ar/ 'back, behind' ∾*ör* /ör/ 'depth'; *xal* /xal/ 'physical suffering' ∾*xöl* /xöl/ 'foot, leg'; *taana* /taan/ 'a wild leek' ∾ *töönö* /töön/ 'a caustic, a cautery'.

6: (a/ü) *ad* /ad/ 'a fiend, devil' ∾*üd* /üd/ 'twelve o'clock noon'; *šar* /šar/ 'yellow' ∾*šür* /šür/ 'coral'.

7: (a/aa) *av* /aw/ 'hunting' ∾ *aav* /aaw/ 'father, papa'; *cana* /can/ 'snow boots' ∾ *caana* /caan/ 'on the opposite bank, the opposite shore'.

20. The phoneme /o/.

The phonological identity of this phoneme results from the following comparisons:

1: (o/I) cf. 18.3.

2: (o/a) cf. 19.2.

3: (o/u) *od* /od/ 'star' ∾ *ud* /ud/ 'weeping willow'; *xor* /xor/ 'poison' ∾ *xur* /xur/ 'rain'.

4: (o/ä) *or* /or/ 'floor' ∾ *er* /är/ 'male'; *orox* /orx/ 'go in' ∾ *erex* /ärx/ 'search'; *bor* /bor/ 'grey' ∾*ber* /bär/ 'bride'.

5: (o/ö) *on* /oŋ/ 'the calendar year' ∾ *ön* /öŋ/ 'fertile (of land)'; *olon* /oln/ 'plenty, much' ∾ *ölön* /öln/ 'hungry';

ono /on/ 'hit it! (imp.)' ∾ *önö* /ön/ 'cold, chill'; *bos* /bos/ 'rise up! (imp.)' ∾ *bös* /bös/ 'a cloth, spread'.

6: (o/ü) *or* /or/ 'floor' ∾*ür* /ür/ 'seed'; *od* /od/ 'star' ∾ *üd* /üd/ 'hugh noon'; *tor* /tor/ 'net' ∾ *tür* /tür/ 'temporarily, a moment'.

7: (o/oo) *odon* /odŋ/ 'award, decoration' ∾*oodon* /oodŋ/ 'short'; *xol* /xol/ 'close, near' ∾*xool* /xool/ 'prepared food'.

21. The phoneme /u/.

The phonological identity of this phoneme results from the following comparisons:

1: (u/I) cf. 18.4.

2: (u/a) cf. 19.3.

3: (u/o) cf. 20.3.

4: (u/ä) *us* /us/ 'water'∾*es* /äs/ 'a cell (of the body or of a honey comb)'; *gur* /gur/ 'male roebuck'∾ *ger* /gär/ 'Mongol yurt'.

5: (u/ö) *ud* /ud/ 'weeping willow'∾*öd* /öd/ 'wing'; *sur* /sur/ 'leather strap, thong'∾*sör* /sör/ 'small prop or support of wood'; *dur* /dur/ 'desire, lust'∾*dör* /dör/ 'nose ring and halter of a cow (different word used for corresponding harness for a camel)'.

6: (u/ü) *ud* /ud/ 'weeping willow'∾*üd* /üd/ 'high noon'; *dur* /dur/ 'lust, desire'∾*dür* /dür/ 'external appearance'.

7: (u/uu) *ul* /ul/ 'sole of the foot'∾*uul* /uul/ 'mountain'; *tul* /tul/ 'for the sake of'∾*Tuul* /tuul/ 'the Tuul River'.

22. The phoneme /ä/.

The phonological identity of this phoneme results from the following comparisons:

1: (ä/I) cf. 18.5.

2: (ä/a) cf. 19.4.

3: (ä/o) cf. 20.4.

4: (ä/u) cf. 21.4.

5: (ä/ö) *er* /är/ 'male'∾*ör* /ör/ 'debt'; *ber* /bär/ 'bride'∾*bör* /bör/ 'entirely, wholly'; *ever* /äwr/ 'horn'∾*övör* /öwr/ 'the bosom of a garment (used as a pocket)'.

6: (ä/ü) *es* /äs/ 'cell (of body)'∾*üs* /üs/ 'hair'; *ter* /tär/ 'that' (dem. adj.)∾*tür* /tür/ 'for a moment'.

7: (ä/ää) *ex* /äx/ 'mother'∾*eex* /ääx/ 'to warm at the fire'; *der* /där/ 'pillow'∾*deer* /dääг/ 'above'.

23. The phoneme /ö/.

The phonological identity of this phoneme results from the following comparisons:

1: (ö/I) cf. 18.6.
2: (ö/a) cf. 19.5.
3: (ö/o) cf. 20.5.
4: (ö/u) cf. 21.5.
5: (ö/ä) cf. 22.5.
6: (ö/ü) *ös* /ös/ 'a grudge, vengeance'∾*üs* /üs/ 'hair'; *dör* /dör/ 'a cow's nose ring and halter'∾*dür* /dür/ 'external appearance'.
7: (ö/öö) *ör* /ör/ 'debt'∾*öör* /öör/ 'other, different'; *xöx* /xöx/ 'blue'∾*xööx* /xööx/ 'to drive, pursue, chase'.

24. The phoneme /ü/.

The phonological identity of this phoneme results from the following comparisons:

1: (ü/I) cf. 18.7.
2: (ü/a) cf. 19.6.
3: (ü/o) cf. 20.6.
4: (ü/u) cf. 21.6.
5: (ü/ä) cf. 22.6.
6: (ü/ö) cf. 23.6.
7: (ü/üü) üd /üd/ 'high noon'∾üüd /üüd/ 'gate'; šür /šür/ 'coral'∾šüür /šüür/ 'broom'.

VOCALIC LENGTH

The opposition of long and short vowels is neutralized in non-initial syllables except in the case of the /i/ phoneme, where its retention is not surprising in view of the occurrence of /i/ as a diphthongal element. In open monosyllables the situation is somewhat unclear: for /i/ we have already cited the pair /či/∾/čii/; for the other phonemes minimal pairs are not to be found. Nevertheless, M. M. H. is able to distinguish long and short in isolation, and such pairs as /ta/ 'you (honorific)' and /saa/ 'milk' are easy to find. In general, though, these pairs do not commute in their syntactic functions, and it would seem that for the most part pronouns and other elements which do not receive prominence

in the sentence have a short vowel while words loaded with an important lexical meaning and thus likely to be given prominence in the chain have a long vowel. The emphatic pronunciation of /či/ and /ta/ with lengthened vowel is also certainly possible. It would seem, therefore, that, although there is no question of neutralization of the distinction of vocalic length in open monosyllables, since it is maintained by the canonical forms of the words, nevertheless the opposition of length may be said to have a minimal functional load in these cases.

PHONETIC REALIZATION OF THE PHONEMES

The following remarks are offered only as approximate indications of the phonetic values assumed by the different phonemes. All consonants and the vowels /a/, /o/, /u/ in initial syllables (but not their long counterparts) display palatalized variants when followed in the same word by non-diphthongal /i/ (but not by /j/) or /ii/.

I. *The Consonants*

1. /b/. This phoneme is articulated as a bilabial unaspirated occlusive, usually unvoiced (IPA [b̥]), especially in word initial position, but often voiced after /m/ or a liquid. In keeping with the general weakness of lip articulation in Mongolian, to be seen also in the vowels, /b/ gives an impression similar to the lenis labial occlusive in English and many other languages; nevertheless, it is the strongest phoneme in the Mongolian labial series.

2. /m/. This phoneme is articulated as a nasal sonant with bilabial closure (IPA[m]).

3. /w/. This phoneme is articulated as a voiced bilabial fricative (IPA[β]), the voiced spirant counterpart of /b/. It is not, however, strictly a continuant, but is formed with dynamic closing and opening motions reminiscent of occlusives, even though complete closure is not achieved: such words as /jawn/ 'go' often sound something like [jaᵘʷ βʷᵘna].

4. /t/. This phoneme is articulated as an aspirate unvoiced fortis occlusive with apico-gingival (i.e. post-dental, rather than alveolar) closure (IPA[t]).

5. /d/. This phoneme is articulated as an unaspirated lenis occlusive with apico-gingival closure, usually voiced (IPA[d]), except in word initial and final position (IPA[d̥]).

6. /n/. This phoneme is articulated as a nasal sonant with apico-gingival closure (IPA[n]).

7. /s/. This phoneme is articulated as a voiceless hissing sibilant (IPA[s]).

8. /š/. This phoneme is articulated as a voiceless hushing sibilant [IPA[ʃ]).

9. /c/. This phoneme is articulated as a voiceless fortis apical hissing affricate. (IPA[ts]). The point of closure is approximately the same as for /t/ and /d/.

10. /z/. This phoneme is articulated as a voiced lenis apical hissing affricate (IPA[dz]), unvoiced in word initial position (IPA[d̥z]). The point of closure is approximately the same as for /c/, /t/, and //d.

11. /č/. This phoneme is articulated as a voiceless apical fortis hushing affricate (IPA[tʃ]). The point of closure for /č/ as well as for /ǰ/ is somewhat more retracted than for /c/ and /z/.

12. /ǰ/. This phoneme is articulated as a voiced apical lenis hushing affricate. (IPA[dʒ]), unvoiced in word initial position (IPA[d̥ʒ̥]).

13. /g/. In word initial or final positions and in contact with voiceless consonants, this phoneme is a voiceless dorso-velar lenis occlusive (IPA[g̊]); intervocalically and in contact with the liquids /l/ and /r/, it becomes a voiced dorso-velar lenis fricative (IPA[ɣ]).

14. /ŋ/. This phoneme is articulated as a dorso-velar occlusive nasal sonant (IPA[ŋ]).

15. /x/. This phoneme is articulated as a voiceless fortis dorso-velar fricative (IPA[x]). It is produced with a very energetic articulation so that such words as /äx/ 'mother' (as Poppe notes, p. 25) give the impression of containing a geminate consonant: [ɛ̈xxə].

16. /l/. This phoneme is articulated as a lateral alveolar liquid (IPA[l]).

17. /r/. This phoneme is articulated as an apicoalveolar fricative flap or light roll. It seems not too different from S. British fricative *r*, and is much less sharp in acoustical effect than the Japanese *r* phoneme. Apparently different from the corresponding Xalxa phoneme, which Poppe describes as strongly rolled (p. 29). (IPA[r̝] or [ɹ]??).

18. /j/. This phoneme is a semi-vowel articulated in the same position as the /i/ phoneme. Sometimes there is sufficient closure for spirant friction to be heard. Where opposed to the phoneme /i/, i.e. in initial postconsonantal prevocalic position, it is shorter in duration and lighter. (IPA[j]).

II. *The Vowels*

19. /a/. This phoneme is an unrounded maximum aperture vowel which is usually similar to the [a] in English *father*. It has, however, a front variant, ranging from IPA[æ] to [ɛ], realized before /i/ or /ii/ in a following syllable.

20. /o/. This phoneme is a back semi-open vowel similar to IPA[ɔ], sometimes rounded, but no more so than in the usual Mid-west American pronunciation of *autumn*. Like /a/, it has a conspicuous front variant before syllabic /i/ or /ii/, which gives an impression varying from that of [ʌ] in *cut* to that of a somewhat retracted [ø]. Lip rounding is definitely more pronounced in the front variant.

21. /u/. This phoneme is realized as a back semi-close vowel, weakly rounded, and approximately IPA[o̜] in phonetic value. When palatalized, it assumes roughly the value of the vowel in South. British *wool* or *wood*, i.e. IPA[ʊ]. With /u/, too, lip rounding is stronger in the front variant.

22. /ä/. A retracted mid vowel, varying in value from a slightly retracted [ɛ̂] to the [ɐ] in the common English pronunciation of *bird*.

23. /ö/. This is a more or less central semi-open vowel with very weak rounding varying from [œ] to [ʌ] in realization.

24. /ü/. This phoneme is a central close vowel with very little lip rounding. M. M. H. equates this vowel in his usage with the unrounded Japanese *u* phoneme. Approximately IPA [ɯᵘ].

25. /i/. This is a close front unrounded vowel varying from IPA [i] to [e] or close [e̝]. Especially when long in word final position or before /ŋ/, the articulation of this phoneme tends to become very open, e.g. *argyn* /argiiŋ/ 'of the means' (gen. case). It will be noted that the Cyrillic orthography employs two graphic forms, *y* and *ij*, for what are identical sounds in the usage being described (cf. *ajany* /ajnii/ 'traveling' in *ajany xün* /ajnii xüŋ/ 'a man traveling, a traveler'; *xünij* /xünii/ 'of the man'; indeed, *ajany xünij* /ajnii xünii/ 'of the traveler'.

It would be of great interest to determine the exact tongue positions assumed in the normal articulation of these vowel phonemes. Although, unfortunately, it was not possible to conduct an x-ray investigation to obtain these data, we were able by comparison with published x-ray photographs by Chiba and Kajiyama of the Japanese and German vowels[9] to form a rather satisfactory estimate of what the approximate average position of articulation of each of the Mongolian vowels must be. These are shown in Diagram 1, the top line representing the contour outline of the top of the mouth, and the lower line, the outline of the tongue in position for the vowel of maximum aperture. The advanced position of /ü/ may seem somewhat surprising; however, we are convinced that its articulation is quite similar to the Japanese *u*, which is shown by Chiba's photographs[10] to be in this position. In view of the wedge-like assymetry of the vocal tract, and the back crowding obvious in Mongolian, this advanced position becomes less surprising.

In Diagram 1 the positions of the palatalized /a, o, u/ phonemes have not been indicated.

The Mongolian vowel configuration may be represented with less physical realism but more clarity on Daniel Jones's vowel quadrilateral as in Diagram 2 (the ranges circumscribed including the palatalized variants).

DISTRIBUTION

I. Vocalic Complexes.

a. Initial Syllables.

In word initial or absolute position, the following syllabic vowels or vowel complexes may occur: /a, o, u, ä, ö, ü, i, aa, oo, uu, ää, öö, üü, ii, ai, oi, ui, äi, öi, üi/.

All of the above may occur after /I/ (which we write herein with *j*) in initial position or after either of the distinct members of the /I/ archiphoneme, /j/ and /i/, in initial syllables after a consonant.

After an initial consonant the combinations /ua/ and /uai/ also occur, though mainly in words of Chinese origin.

[9] Cf. T. Chiba : *Japanese Vowels*, 1931, pp. 27-28, and T. Chiba and M. Kajiyama : *The Vowel, Its nature and Structure*, 1941, p. 150.

[10] *J. Vowels*, p. 30

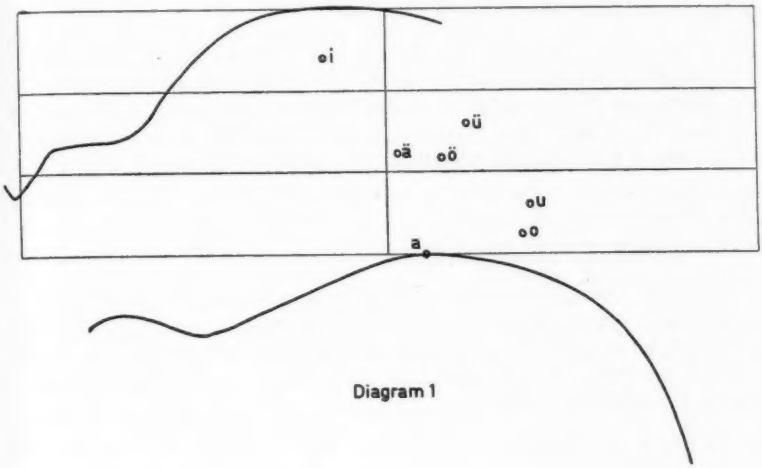


Diagram 1

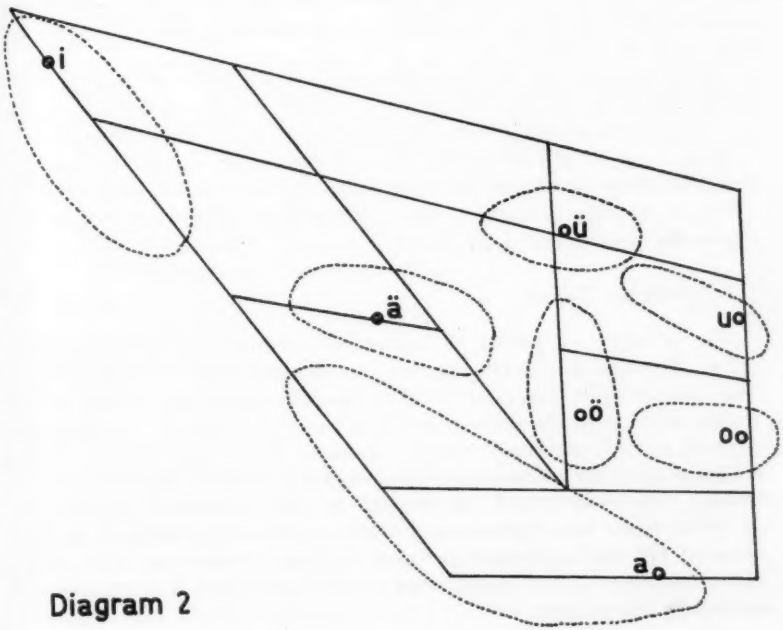


Diagram 2

b. Non-initial Syllables.

In syllables following the first, the following DISTINCTIVE vowels or vowel complexes occur: /a, o, u, ä, ö, ü, i, ii, ai, oi, ui, äi, öi, üi/.

There are also non-initial syllables in which none of the above occur. It is therefore impossible in Mongolian to determine the number of syllables by counting the number of vowel phonemes.

II. Consonants.

a. Word and Syllabic Initials.

In both word and syllabic initial positions the following consonants may occur: /b, m, t, d, n, c, z, č, ǰ, g, w, s, š, x, l, r, j/. Except for combinations with /j/, no initial clusters occur. In word initial position the phonemes /r/ and /w/ are somewhat rare, but must be reckoned with.

b. Syllabic Finals.

Except for the reservation that /n/ does not occur before an initial /b/ of a following syllable (/n/ → /m/), the following consonants may occur in syllable final position: /m, t, d, n, c, z, č, ǰ, g, ŋ, w, s, š, x, l, r/.

Final clusters are also formed from a sequence of one of the group /m, n, g, l, w, r/ followed by one of the group /t, d, c, z, č, ǰ, s, š, x/. In addition, the clusters /st/ and /xt/ also occur.

Where in a sequence the same consonant occurs twice, as syllabic final and adjacent syllabic initial, both are clearly pronounced, the acoustical effect being that of a long geminate, e.g. *sallax* /sallx/ 'to float a raft'.

III. Syllabic Types.

The number of peaks of acoustical intensity in a Mongolian word often varies with the speaking tempo and other circumstances. The phonetically realized syllable therefore has no determinate stable relation to the distinctive structure of the word. In normal current speech startlingly long sequences of consonants are often realized with no more than an occasional vocalic murmur intervening. In slow speech, on the other hand, consonant sequences of more than two phonemes or sequences not permitted in the rules of IIb are separated by weak vowels of indeterminate color harmonizing with the realizations of the distinctive (i.e. phonemic) vowels in the words.

For the purposes of this paper, we shall define a (phonological) syllable (as opposed to a phonetic syllable) as a member of one of a limited number of phoneme sequence types which form integral fractions of words. Every word must be either of one syllable or of an integral number of syllables, and every sequential phoneme must be definitely assignable to one particular syllable. In this sense, the syllable is a purely operational concept and a different, though equally valid, classification of sequence types (syllables) might require the assignment of a particular phoneme in a particular sequence to a different syllable than would be the case in terms of the present analysis. In some languages, (e.g. English) this concept of syllable would be useless because of the impossibility of deciding in many cases to which of two syllables a given phoneme belongs (e.g. in *cabin* the /b/ cannot be said to belong more to the first than to the second syllable).

All Mongolian syllables may be regarded, therefore, as combinations of the following type elements:

1) Vowel complexes (V) of the following types
   a) a simple vowel
   b) a long vowel
   c) a diphthong, i. e. simple vowel plus /I/.
   d) /i/ or /j/ or their archiphoneme followed by a simple vowel, a long vowel, or a diphthong.
   e) /ua/ or /uai/ in words of Chinese origin.
2) The consonants /m, n, g, l, w, r/, which we shall represent as a group with the letter *C*.
3) The consonants /t, d, c, z, č, ǰ, s, š, x/, which we shall represent collectively with the letter *K*.
4) /b/
5) /ŋ/
6) /j/
7) /st/
8) /xt/

In terms of these eight types of elements, we may describe the structure of any Mongolian syllable as representing one of two alternate formulas, either (C/K/b/no consonant)+V+(C/K/CK/ŋ/st/xt/no consonant) or (C/K/b/j)+no vowel+(C/K/CK/ŋ/st/xt).

The second formula is not represented in any initial syllable. The phoneme /j/ does appear in the first formula since it is included

in the sign for the vowel complex. As indicated in Section I (Vocalic Complexes), the possible vowel distinctions subsumed under the vocalic complex are different in the initial and non-initial syllables.

There is a strong dissimilative tendency, which might be taken note of, for lenis consonants to become fortis before another lenis in suffixation, e.g. *baɬdax*/batd.x/ 'to become hard' (intrans. inf.), *baɬaɬgax* /batt.gx/ 'to make hard' (trans. inf.). However, this tendency has not yet gone far enough to be regarded as a phonological prohibition.

IV. Phonetic Syllables.

All syllables containing a phonemic vowel are realized as phonetic syllables with the exception of non-initial syllables containing only a simple /i/, which may not be heard in current speech though its characteristics are preserved in the palatalization of the preceding consonants and vowels.

Syllables without phonemic vowels will usually contain a neutral vowel which will be more front or back, open or close, according to the character of the phonemic vowel in the initial syllable. This neutral vowel is, as is to be expected, somewhat unstable and often disappears in very rapid speech.

In extremely deliberate speech every consonant except /ŋ/ may be supported by a neutral vowel sound following it, resulting in a phonetic sequence of open syllables. It is, however, essential to the character of the phoneme /ŋ/ that it *not* be followed by any vocalization before the following consonant onset or hiatus.

Especially after a voiceless final consonant, there is a tendency for a reduced vowel to be produced.

As a rule, in a sequence CVCCC, the reduced vowel will occur between the last two consonants eliminating all consonant clusters.

It often happens accordingly that the syllabic structure, both phonetic and phonological, of a root will alter with suffixation, e.g. *iteg* /i-tg/ [itəg] 'believe' (imp. or stem), *itgex* /it-gx/ [itgəx] 'to believe' (inf.). However, the reduced vowels being unstable, cases of phonetic metatheses are common, especially in contact with /r/ and /l/. Poppe[11], using a different transcription, cites the examples /arg/ [arga] [arag] 'means', /xaalg/ [xa:lga] [xa:lag] 'gate', /xaalgtai/ [xa:lgatai] [xa:lagtai] 'with a gate', /terg/ [terge]

[11] *Op. cit.* p. 12.

[tereg] 'cart'. To this we might add, that, as a consequence perhaps, the location of the reduced vowels written in the Cyrillic orthography is often quite arbitrary (*sallax* /sallx/ but *ilelt* /illt/).

V. Vowel Harmony.

There is a general tendency for native words in Mongolian to follow a rule which restricts the vowels in nonfirst syllables to those which accord in certain characteristics with the vowel in the first syllable. The existence of exceptions to this rule preclude our regarding it as properly a principle of Mongolian phonology; it rather belongs to the morphophonology of the word. Were there no exceptions we might describe Mongolian with only four vowels, i. e. /a, o, u, i/, plus two opposed characteristics (front∽back) appertaining not to the phoneme segment but to the word. The rules of vowel harmony hold sufficiently well, as it is, for words violating these rules to be definitely conspicuous. In native words, a break in vowel harmony within the word almost invariably marks a morphemic boundary (though such a break is not phonologically predictable). Except for words produced with a number of extremely frequently occurring suffixes which do not change for vowel harmony, words contradicting vowel harmony are mostly of exotic origin. These occur so commonly, however, that they cannot be neglected. Nevertheless, the rules of vowel harmony have a genuine validity as representing the prevailing vowel structure of native morphemes, if not altogether that of native words, and as such are here presented.

The vowels may be divided into two groups which do not occur together in the same morpheme: a back group (called 'masculine' by native grammarians) /a, o, u/, and a front group ('feminine') /ä, ö, ü/. The phoneme /i/ may occur as a following vowel in words with either front or back vowels; however, words containing no vowels other than /i/ (or /ii/, of course) take front vowel suffixes. Foreign loan words with /i/ in the initial syllable followed by a back vowel, almost invariably change /i/ to /ja/ or /a/ in the spoken language, e.g. *dijan* /dijaŋ/→*dajan* /dajaŋ/. M. M. H. regards, indeed, the entry *ildam* /ildm/ in Rinčine's dictionary[12] as a positively mistaken entry for what could only stand under the rubric *jaldam* /jaldm/.

[12] A. R. Rinčine : *Kratkij Mongol'cko-russkij slovar'*, Moscow, 1947. p. 118

The mid-open vowels /o/ and /ö/ can be followed immediately in the next syllable only by (respectively) /u, o, oi, i, ii/ or /ü, ö, öi, i, ii/. If /u/ or /ü/ follows its corresponding mid-vowel, then /a/ or /ä/ may appear in the third syllable. In bisyllables, the sequence is restricted even more, so that after /o, oo, oi/ or /ö, öö, öi/ in the first syllable, only /o, oi, i, ii/ or /ö, öi, i, ii/ may occur.

These rules we may depict schematically as follows:

A. In bisyllabic words.

Back Vowels

1st syllable	2nd syllable	
a aa ai u uu ui	a u i	ai ui ii
o oo oi	o i	oi ii

Front Vowels

1st syllable	2nd syllable	
ä ää äi ü üü üi	ä ü i	äi üi ii
ö öö öi	ö i	öi ii

B. In polysyllabic words.

In the following tables, the left hand marginal column lists the vocalics occurring in initial syllables, while the top marginal row lists all vocalics (as already given in table A) which occur in second syllables. The non-occurrence of a particular second syllable vocalic after a given first syllable vocalic is indicated by a diagonal cross in the space representing the intersection of the rows and columns of the respective vocalics. The vocalics noted in the spaces are those which may occur in the third syllable of a word in which the first and second syllable vocalics correspond with those determining the space in the table. For instance, if the first syllable vowel is /a, aa, ai, u, uu/ or /ui/ and the second syllable vowel is /a, ai, u, ui, i/ or /ii/, then there may follow in the third syllable /a, ai, u, ui, i/ or /ii/. If the first syllable vowel is /o, oo/ or /oi/ then /a/ may occur in the third syllable only if the second syllable vowel if /u/.

In words of more than three syllables, the vowel harmony excludes sequences of three vowels, at any place in the word, which contradict this pattern.

Back Vowels

	a ai	u	ui	i ii	o oi
a aa ai u uu ui	a ai		u ui	i ii	
o oo oi		a i u ii		o oi	u i ii

Front Vowels

	ă ăi	ŭ	ŭi	i ii	ö öi
ă ăă ăi ŭ ŭŭ ŭi		ă ăi	ŭ ŭi	i ii	
ö öö öi		ă i ŭ ii		ö i ŭ öi ii	
i ii	ă ăi		ŭ ŭi	i ii	

ANALYSIS AND CLASSIFICATION OF THE PHONEMES

A. The Consonants.

In considering the Mongolian consonant pattern, we may simplify our task by immediately setting aside the liquids /l/ and /r/ and the semi-vowel /j/. These fall outside the main consonant pattern; /l/ and /r/ together form one group opposed to all the other consonants by their intrinsic characteristics as liquids, and /j/, being a vowel in consonant function, is uniquely opposed to the other consonants by its vocalic nature.

The sibilants /s/ and /š/ and the affricates /c, z, č/ form a clear-cut and symmetrically organized system built upon the oppositions *sibilant∽affricate*, *hissing∽hushing*, and *fortis∽ lenis*.

		hissing	hushing
sibilant		s	š
affricate	fortis	c	č
	lenis	z	ǰ

The remaining consonants are easily divided into three orders, which we may call *labial* (/b, w, and m/), *apical* (/t, d, and n/), and *dorsal* (/x, g, and ŋ/). The phonemes /m, n, and ŋ/ form a nasal correlation with the other members of each order. The non-nasal dentals /t, d/ are clearly related in the same way as the members of each of the two affricate series, i.e. as *fortis* versus *lenis*.

In considering the labial series, our first impulse might be to assign /b/ with its phonetically weak articulation to the position of labial lenis stop, opposed to a continuant /w/, and leaving a hole in the pattern at the labial fortis position. It is worthwhile observing that, although there can be no doubt that the opposition of /b/∞/w/ is firmly established by well naturalized minimal pairs, it is not heavily loaded functionally, and if all loan words—even those of long standing—were excluded from the lexicon, [β] would appear simply as the weak positional variant of the phoneme otherwise realized as [b̥] or [b]. There can be no doubt about the fact that /b/, despite its weak articulation, is strong with respect to /w/. With only two oral phonemes in the labial order there would seem to be no reason to regard this distinction of strong and weak articulation as other than that which we have already characterized in the apical order and in the affricates as *fortis* versus *lenis*. Any doubt we might have about including /w/ in a correlation in which occlusive articulation predominates should be dispelled by consideration of the dynamic articulation of this phoneme, often with rather clean cut closing and opening phases, more typical of occlusive than of continuant articulation.

In the dorsal series, we see a vigorously articulated fortis /x/ (so strong that it sometimes gives the acoustical impression of gemination) opposed to a weak /g/ which actually loses its occlusion in weak environments to become [ɣ].

Remarkable though it may appear, we thus find ourselves with a fortis-lenis correlation of /b, t, x/ versus /w, d, g/. Neither voice

nor aspiration plays a part in this opposition, and the group as a whole is not characterized by occlusive as opposed to continuant articulation, since /w/ and /x/ are never occluded and /g/ is realized as a fricative in about half of its occurrences. Nevertheless, /b, t, x, w, d, g/ and their corresponding nasals are characterized by a sort of dynamic intermittentness which makes them less capable of sustained production than /s, š, r, l, j/ or the affricates (whose sibilant releases can be prolonged). We may therefore summarize the system of the intermittent consonants as follows:

		Labial	Apical	Dorsal
oral	fortis	b	t	x
	lenis	w	d	g
nasal		m	n	ŋ

The system of affricates may be connected with the pattern of intermittent consonants as two added dimensions of the apical order.

B. The Vowels.

One is strongly tempted to seek a phonological classification of the Mongolian vowels in terms of distinctive features which will reflect or explain the distributional tendencies of the Mongolian vowel harmony. The notation we have used in this paper, in which the 'front vowels' are represented by the same signs as the 'back vowels' with the diaeresis added, probably encourages this inclination. If this be so, it is a fault of the notation, since no matter how Mongolian may have been predisposed to the development of the present pattern of vowel harmony by its phonological structure in an earlier stage, there is little reason to believe that the preservation of this vowel harmony is due to anything except the conservatism of lexical forms. In the first place, lip rounding is extremely weak—as weak as in Mid-western American English—as is shown by M. M. H.'s personal recognition of the resemblance of his /ü/ phoneme with the Japanese (completely unrounded) *ɯ̈* phoneme. Indeed, lip rounding is probably only

an accessory feature brought into play to reinforce close distinctions. It is indicative that /o /and /u/ are much more clearly rounded when shifted forward by palatalization and therefore in danger of being confused with /ö/ and /ü/ which resemble them but have even weaker lip rounding. Even the *front/back* contrast determined by the vowel harmony cannot be said to represent a phonological opposition, as shown by the fact that /a/, when under the palatalizing influence of a subsequent /i/, may well be articulated as far forward as /ä/ or /ö/. The articulatory data show clearly, as we should have expected from the general exception formed by loan words, that the vocalic pattern is structured without specific regard to vowel harmony. We have, rather, to do with a triangular system of three classes with three degrees of aperture which we might diagram thus:

	front	central	back
close	i	ü	u
mid	ä	ö	o
open		a	

Position is probably not alone distinctive but is combined with rounding so that the front class is both front in position and unrounded, the back class is both back in position and rounded in respect to the much more weakly rounded central class, which nevertheless partakes of rounding enough that it can serve to resolve close resemblances between /ä/ and /ö/.

B. The Phonemes as Bundles of Relevant Features

In the foregoing we have taken for granted a basic difference between vowels and consonants. It would not be accurate, however, to regard this difference as one of opposition by commutation: generally the vowels function as syllabic nuclei while the consonants function as the determinants of syllabic nuclei. Except in the case of /j/, it cannot be said that the vowels and consonants commute; they are not opposed in the pattern, but rather contrasted in the *chain*. The *capacity* for opposition, however, is retained in the respective characters of consonants and vowels, which differ basically in accordance with their contrasted functions, and is indeed realized in the case of the semi-vowel /j/, which is a vocalic phoneme (appearing in the vowel pattern as /i/) in consonant function.

The vocalic phonemes have each been further characterized as having one of three grades of aperture (*open*, *mid*, or *close*) and belonging to one of three positional series (*front*, *central*, or *back*), distinctions of position being lost in the degree of maximum aperture.

The consonant phonemes, besides /j/, may be characterized as *intermittent* or *continuant* or *affricate* according to the type of articulation. Within the continuants we find a further opposition of *liquid* versus *sibilant*. The affricates are intermittent-sibilant complexes and may be connected with either the intermittent or sibilant patterns. The liquids are either *r-type* or *l-type* phonemes. The sibilants and affricates are of two types: *hissing* and *hushing*. The intermittent phonemes each belong to one of three homo-local series (*labial*, *apical*, or *dorsal*) and form one correlation of *nasal* versus *oral* features and again, within the oral phonemes, a *fortis* vs. *lenis* correlation. The affricates also enter into the *fortis-lenis* correlation.

We arrive, therefore, at the following inventory of relevant features:

1) *vocalicity* vs. *consonantality* (Opposed only when the functional contrast between vowels and consonants is lost in the case of /j/.)

Combining with vocalic features:

2) *open* vs. *mid* vs. *close* aperture
3) *front* vs. *central* vs. *back* position

Combining with consonant features:

4) *intermittent* vs. *continuant* vs. *affricate* articulation

Combining with continuant articulation:

5) *liquid* vs. *sibilant* striction

Combining with liquid striction:

6) *r-type* vs. *l-type*

Combining with sibilant striction and affricate articulation:

7) *hissing type* vs. *hushing type*
8) *labial* vs. *apical* vs. *dorsal* orders
9) *nasality* vs. *orality*

Combining with intermittent oral and affricate articulation:

10) *fortis* vs. *lenis* articulation

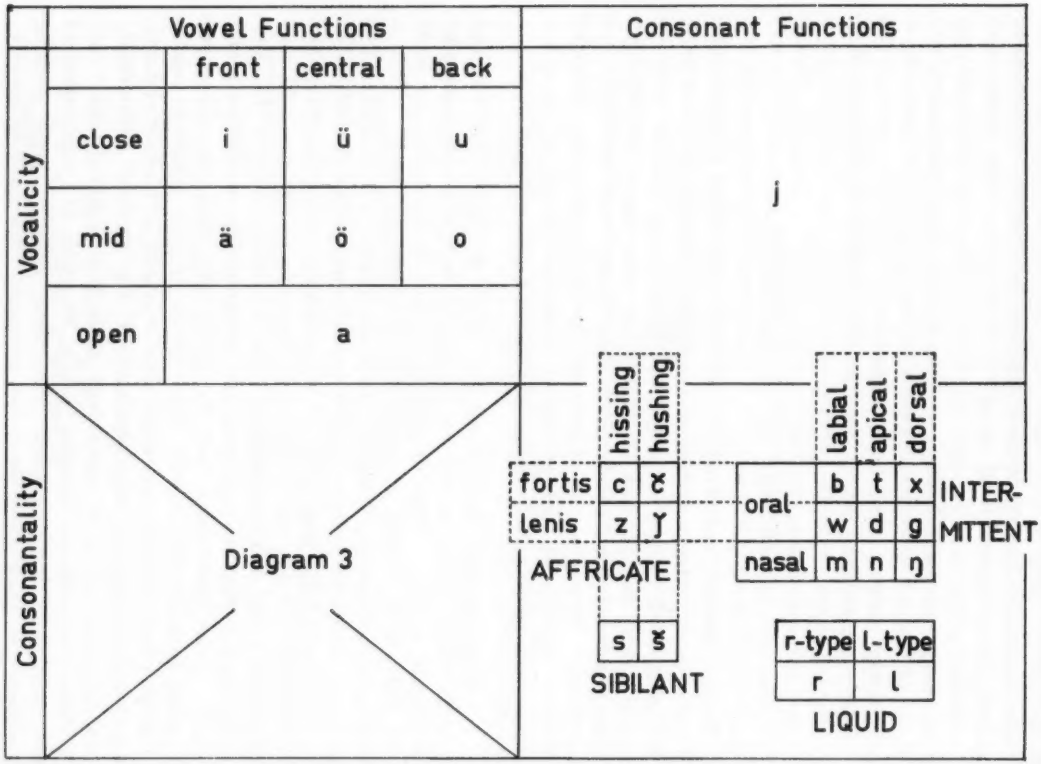
No doubt most of the above listed relevant features are phonetically very complex. It may be that further analysis would make it possible to reduce these features to elements which would permit the recognition of identical features in both vowel and consonant patterns. To date attempts in this direction involving other languages have been more stimulating than convincing, and in the absence of instrumental data for Mongolian, further analysis here would only involve us in conjecture.

The complete phoneme system is summarized in Diagram 3.

CONCLUDING REMARKS

Poppe, Austin, and Torjesen all deal with what is essentially for practical purposes the same language as that considered here. It cannot, of course, be expected that theirs should coincide exactly with the present treatment, or even agree among each other, because the specific individual usages are certainly not identical. However, the system set forth in this paper diverges in some respects to such an extent from that of any of the other researchers as to call for some explanation and comment.

1. It is notable that in all the other analyses, the reduced vowels which occur to relieve consonant sequences are accorded phonemic status. This, of course, makes it necessary to treat the distinctive vowels in non-initial syllables as 'long'. This is not so objectionable; however, the consequence which makes this analysis completely unacceptable is that one is then forced either to declare that it is impossible to say precisely what is the phonemic constitution of any particular word because of the free and easy appearance and disappearance of these vowels or, what is the same thing, to say that it is not possible to define the conditions under which these 'phonemes' will be physically realized. The new Cyrillic orthography, which also writes in the reduced vowels (though often naturally in erratic disagreement with other systems, e.g. where Poppe has *jala* 'punishment' the Cyrillic orthography has *jal*; while *jal* is in agreement with Austin, the Cyrillic contradicts him with *ex* 'mother' where he gives the equivalent of *exe*) will doubtlessly burden Mongolian children with spelling problems, as they must learn the arbitrary location and identity of interconsonantal vowels. Of course, it would be possible to distribute these interconsonantal vocables according to set rules


Vowel Functions
Consonant Functions
front
central
back
close
i
ü
u
mid
ä
ö
o
open
a
j
Vocalicity
Consonantality
hissing
hushing
labial
apical
dorsal
fortis
c
ƈ
lenis
z
ʒ
AFFRICATE
oral
b
t
x
INTER-
w
d
g
MITTENT
nasal
m
n
ŋ
s
ʂ
SIBILANT
r-type
l-type
r
l
LIQUID
Diagram 3

in relation to the consonantal structure of each word,[13] but this is to say that the occurrence of these vowels is predictable and that they can serve no distinctive function. Which is the case.

Hattori[14] admits a solution like the one presented in this paper as possible but objects to it on the grounds (1) that it is motivated by the desire for a concise notation which is not the immediate objective of phonology, and (2) that phonology must make clear the mutual relationships of the socially conventional forms in the pronunciation of the speakers of a language and the structures which appear within these relationships. It is not altogether clear to us what Hattori has in mind here. It is of course true that this solution does result in a more concise notation, but this is incidental to the fact that it represents truthfully the phonologically distinctive elements and *nothing more*. How this fails to accord with Hattori's second requirement of phonology we do not understand.

It is worth remarking that, as can be seen from many examples given in the body of this paper, the Cyrillic orthography makes use of the fact that /n/ and /ŋ/ are distinguished only in absolute word final position to make a graphic economy by writing both with the same sign and distinguishing them by adding the reduced vowel (which can only be heard in very slow speech) after /n/.

2. Austen and Torjesen, as summarized in Hattori's article, while treating the interconsonantal vocables as phonemes, surprisingly discard, in some words, the short /i/ in the non-initial syllable and account for the semantic difference between such pairs as *al* /al/ 'kill.' (imp.)~*al*^{b} /ali/ 'give me!' (uninflected imperative ejaculation) by positing three extra palatalized vowel phonemes to oppose each of the short vowels /a/, /o/, and /u/, which do indeed have palatalized variants before /i/. Thus Austin writes e.g. *mór* for 'horse', whereas we should transcribe /mori/. We cannot accept this as a possible analysis for the usage of M. M. H. It is

[13] To a certain extent the modern Cyrillic orthography does this. Though greatly complicating the spelling, the writing in of the non-distinctive reduced vowels does not obscure the pronunciation, if the practice of writing the phonologically distinctive vowels as long or double is consistently followed Unfortunately, the modern orthography is not altogether consistent: some of the exceptional spellings of native words can be explained, e.g. *morior* /morior/ and *bariax* /bariax/ which are so written in conformity with an orthographic rule prohibiting the writing of three vowel signs together and in which it is clear from the position that a true distinctive vowel is intended; the pronunciation of such loan words as *kapital* and *kilogram* is not clear from the spelling.

[14] *Op. cit.* p. 101.

true that the /i/ may not be very clearly heard in current speech; however, it is pronounced no more clearly after a syllable with a long vowel, which does not display a palatalized variant than after a short vowel, and to assert the presence of the /i/ phoneme in the one case and to ascribe its distinctive function to the palatalized vowel variant in the other seems to us a highly artificial and arbitrary solution. Moreover, the occurrence of palatalized variants of the consonants before /i/ must also be reckoned with. The following examples might be cited: *Bajgaal* /baigal/ (dialect pron.) 'Lake Baikal' [baiγal]∞*bajgalᵇ* /baigali/ 'nature' [baiγal'[1]] or [baiγal']; *Tuul* /tuul/ 'the Tuul River' [tọ:l]∞ *tuulᵇ* /tuuli/ 'fairy tale' [tọ:l'[1]] or (tọ:l'].

These examples would suggest the possibility of dropping non-initial short /i/ in such environments as do not insure its full realization, transferring the differentiative responsibility to the preceding palatalized consonant variant. This would entail the augmentation of the consonant system with a complete set of palatalized consonant phonemes opposed to all of the unpalatalized consonants which can occur in syllable final position. Aside from the fact that this is poor economy, we cannot agree to such a solution for various reasons. For example, there is the fact that the phoneme /i/ can occur before any consonant that can function in post-vocalic position. This makes it difficult to regard the distinctive characteristics of /i/ as always phonologically simultaneous with those of one of the palatalized consonants since the sequence [/i/-*characteristics followed by consonant characteristics*] must be opposed either to [*consonant characteristics* with /i/ *characteristics followed by* /i/ *characteristics*] or to [*consonant characteristics with* /i/-*characteristics*], the two latter sequences being free variants.

3. The Cyrillic orthography makes the distinctions *y*∞*ij* and *i*∞*ь*. It would be interesting to know if there is a usage which makes such distinctions in pronunciation or if these are purely orthographical. They are in the usage here described merged so that *y* or *ij* are both /ii/ and *i* and *ь*, both /i/.

4. As has been previously implied, a notation representing only the phoneme inventory of pure Mongolian words and loan words of long standing will not be capable of representing the contemporary speech of educated Mongols, many of whom are polyglots able and inclined to use foreign words without assimilating their pronunciation. Foreign words in the speech of the sophis-

ticated may pass without affecting the basic phonological pattern; yet, a very large number of loan words from Russian and Chinese have gone beyond cultural and literary circles and have become widely current. Among other influences, these loan words have brought strong pressure to bear for the introduction of /p/ and /k/ phonemes. Since there are no holes in the existing pattern of intermittent consonants, this can only be effected through the development of aspiration as a pertinent feature. In spite of the natural resistance which might be expected against this, it wou d seem that already there are so many words with *p* which are well naturalized that it probably must be given phonemic status, if not merely native words and long standing loans, but the total active vocabulary is considered. In the case of *k*, the situation is not clear: there seems to be greater resistance to this innovation and the truly unlearned usually pronounce such recent loans as *kapital* as /xapital/; nevertheless, educated persons such as M. M. H. are certainly able to make a distinction between [kh], [x], and [g]. There are also a certain number of foreign words written with *f*, such as *fabrik*. Polyglot speakers such as M. M. H. pronounce such words, of course, without any difficulty. Most speakers, however, pronounce *paris* 'Paris' and *fabrik* with the same initial sound. An /f/ phoneme would seem to require an opposition of voicing to differentiate it from the /w/ phoneme: the ordinary Mongolian, already struggling to master distinctive aspiration, may not rise to this second demand unless functional loading in the lexicon makes this distinction more important than it is at present.

SAMPLE TEXT

The following is a translation of the story of the Sun and the Wind to be found in *The Principles of the International Phonetic Association* (London, 1949).

/ärt näg cagt nar, salxi xojor xŭčä šalgǰ üzx boljä. tägäd, ajnii xünii daxiini tailulǰ čadsnni diilsŋ boln gäǰi togtod, salxi uridar orldǰ äxlw. salxi xŭčniixä činägär üläŋ äxlsnd, ajnii xüŋ, salxi salxilx tutm daxa čang atgǰ bijdä naaldulw. tägtl narnii äälǰ boljä. täxlär nar üülnii zawsrar amrlingui čaraiga gargǰ, dulaŋ ilčä ilgäsnd, ajnii xünii sätglni ulm ulmar amglŋ bolǰ, süüldni daxa tailčixjä./

THE PHONETIC TREATMENT OF THE ARABIC LOANWORDS IN ETHIOPIC

WOLF LESLAU

There are various reasons for the existence of Arabic loanwords in Ethiopic.[1] Throughout its entire history, as a spoken and a literary language, Geez was exposed to the influence of Arabic. Indeed we know that most of the Ethiopian religious writings were translated from Coptic, Greek, Syriac, and Arabic. As a spoken language, Geez as well as the other Ethiopic languages was in constant contact either with Arabic speakers or with Moslem Ehiopians. There had always been a migration of Arabic speakers into the various regions of Ethiopia. These Arabic speakers found commercial opportunities in Ethiopia and became the merchants of the country.[2] They also founded political entities there,[3] and converted certain segments of the Ethiopian population to Islam.[4] This is particularly the case of the Tigrinya and Tigre speakers in the north, and of the Harari, Gurage and Argobba communities in the south. The Arabs on the whole live dispersed among their Ethiopian neighbors, but there are also concentrated Arab villages having constant relations with the Ethiopians. For all these reasons it is not surprising to find a certain number of Arabic loanwords in Ethiopic. The present study will deal with the phonetic treatment of these loanwords.[5]

[1] The names of the languages are abbreviated as follows: A. = Aymellel, Amh. = Amharic, Ar. = Arabic, Arg. = Argobba, C. = Čaha, E. = Eža, Ed. = Endegeň, En. = Ennemor, G. = Geez, Go. = Gogot, Gt. = Gyeto, Gur. = Gurage, Har. = Harari, M. = Muher, Ms. = Masqan, S. = Selti, Te. = Tigre, Tna. = Tigrinya, W. = Wolane, Z. = Zway

[2] In the Moča dialect of Kafa, the word *nägado*, literally 'merchant' means 'Moslem merchant'.

[3] See especially E. Cerulli, *Studi etiopici I. La Lingua e la storia di Harar*, Rome, 1936, p. 1-56, and bibliography in the footnotes.

[4] J. S. Trimingham, *Islam in Ethiopia*, Oxford. 1952.

[5] For more detailed studies of the loanwords in the various Ethiopic languages, see W. Leslau "Arabic loanwords in Harari", in *Studi orientalistici in onore di Giorgio*

Except for the literary stage of Geez that borrowed Arabic words from classical Arabic, all the spoken languages took over the loanwords from the spoken Arabic dialects. Considering the geographic situation and the mutual relations one could suspect an Egyptian, Sudanese or a South Arabic origin of the various loanwords. Unfortunately, our knowledge of the phonetic system of some of these dialects is still insufficient to allow a definite statement. In many cases we certainly have to consider more than one Arabic dialect as the possible source. For all these reasons I have taken the classical Arabic type as basis for the analysis and compared the loanwords with the source material of classical Arabic.

For the benefit of the reader I should like to mention here briefly the distribution of the Ethiopic languages. The languages of the north are Geez, Tigre, and Tigrinya; the South Ethiopic languages include Amharic, Argobba, Harari, Gafat, and Gurage. The language of Geez is no longer spoken; it remained the language of the liturgy. It is difficult to determine the exact time when Geez ceased to be spoken and was replaced by Amharic. This process might have taken place some time between the 12th and 14th centuries. Tigre is spoken in the northern part of Eritrea, in the region of Keren (up to the border of the Anglo-Egyptian Sudan), in Massawa, and on the islands of Dahlak. Tigrinya is spoken in the northern part of Ethiopia and Eritrea. Amharic is used in the central part of Ethiopia and is the national language of the country. Argobba is spoken in the region of Ankober, north of Addis Ababa; it was also spoken in the region south of Harar. Harari is spoken in the city of Harar. Gafat was spoken in the southern part of Godjam, in the region of the Blue Nile; at present it is spoken only by a few individuals. Gurage is a language cluster spoken in the region of Gurage, situated south-west of Addis Ababa. It is divided into three main groups: a) East Gurage including Selti, Wolane, Ulbarag, Inneqor, and the dialects spoken on the islands of the lake Zway; b) West Gurage including Chaha, Eža, Ennemor, Endegeɲ, Gyeto; with a possible

Levi della Vida, vol. 2, Rome, 1956, p. 14-35; "Arabic loanwords in Tigre", *Word*. 12 (1956). 125-141; "Arabic loanwords in Tigrinya", *JAOS* 16 (1956). 204-213. In print: "Arabic loanwords in Geez", "Arabic loanwords in Amharic", "Arabic loanwords in Gurage", "Arabic loanwords in Argobba" The reader will find in these studies the criteria helping us in the determination of a loanword.

sub-group of Muher, Masqan, and Gogot; c) Northern Gurage with Aymellel as its only representative.

In order to analyze the phonetic treatment of the loanwords we must examine first the phonetic system of Arabic and of the various Ethiopic languages. The consonant phonemes of these languages are as follows.[6]

Here are a few observations on the phonetic system.

ARABIC and ETHIOPIC. We find that with the exception of of p, $\underline{p}$, $\check{c}$, $\underline{\check{c}}$, $\check{z}$, $\check{n}$, of the palatal series g', k', q', x, and of the labialized series g^{w}, k^{w}, q^{w}, $\underline{h}^{w}$, b^{w}, f^{w}, m^{w}, p^{w}, all the other phonemes of Ethiopic are also found in classical Arabic. Arabic, on the other hand, has phonemes that are not found in Ethiopic. They are: the interdentals $\underline{t}$, $\underline{d}$, and the emphatic $\underline{t}$; the lateral (?) emphatic $\underline{d}$; and the velar voiced $\dot{g}$.

The sounds $\underline{t}$, $q(\underline{k})$, $\underline{s}$, $\underline{p}$, $\underline{d}$, and $\underline{\check{c}}$ are in Ethiopic of the "glottalized" type whereas $\underline{t}$, $q(\underline{k})$, $\underline{s}$, $\underline{d}$ and $\underline{t}$ are in Arabic of the "velarized" type.

GEEZ. Since Geez is no longer spoken we do not know the exact pronunciation of certain sounds that no longer exist in the spoken languages. This is the case of the sounds transliterated $\underline{d}$ and $\check{s}$. In the present-day pronunciation of Geez as used by the priests, $\check{s}$ is confused with s, and $\underline{d}$ with $\underline{s}$. Likewise, the laryngeals and $\underline{h}$ are already confused in Geez manuscripts from the 13th century onward.

TIGRINYA. A characteristic feature of Tigrinya is the spirantization of the velars k and q. These velar stops become spirant $\underline{k}$ and $\bar{q}$ when in intervocalic position or after a vowel. Thus, *käfätä* "open", but *yəkäffəl* "he opens"; *qäbärä* "bury", but *yəqäbbər* "he buries"; *bäqli* "mule".

AMHARIC. The sound $\underline{s}$ exists only in North Amharic. In South Amharic it is confused with $\underline{t}$. The glottal stop ' occurs in Amharic as an occasional representative of the glottalized q.

HARARI. The sound $\underline{p}$ is not attested, but it probably exists in isolated examples. The velar voiced $\dot{g}$ and the laryngeal h occur only in Arabic loanwords. The velar spirant $\underline{k}$ is often a phonetic variant of k.

ARGOBBA. The glottal stop ' occurs in Argobba only in Arabic loanwords. The laryngeal $\underline{h}$ exists in Argobba words only as geminated sound, in connection with h; thus *hallăba* "to milk",

[6] Gafat is not treated in this study

	ARABIC	GEEZ	TIGRE	TIGRINYA	AMHARIC	HARARI	ARGOBBA	GURAGE
Labials	*b f* *m w*	*b f* *m w* *p ṗ*	*b f* *m w* *p ṗ*	*b f* *m w* *p ṗ*	*b f* *m w* *p' p.*	*b f* *m w* *(p)*	*b f* *m w* *(p)*	*b f* *m w* *p ṗ*
Dentals and interdentals	*d t* *ḏ ṯ* *ḍ ṭ* *ṭ*	*d t* *ṯ* *ḍ*	*d t* *ṯ*	*d t* *ṯ*	*d t* *ṯ*	*d t* *ṯ*	*d t* *ṯ*	*d t* *ṯ*
Sibilants	*z s* *ṣ*	*z s* *ṣ*	*z s* *ṣ*	*z s* *ṣ*	*z s* *(ṣ)*	*z s*	*z s*	*z s*
Liquids	*l n* *r*	*l n* *r*	*l n* *r*	*l n* *r*	*l n* *r*	*l n* *r*	*l n* *r*	*(l) n* *r*
Velars	*k q* *ḫ ḳ*	*g k* *ḫ* *q*	*g k* *q*	*g k* *q*	*g k* *q*	*g k ḳ* *q* *|ǧ|*	*g ḳ* *q*	*g k (ḳ)* *q*
Labiovelars and rounded		*g ʷ k ʷ* *ḫ ʷ* *q ʷ*		*g ʷ k ʷ* *q ʷ*	*g ʷ k ʷ* *h ʷ* *q ʷ*		*g ʷ k ʷ* *q ʷ*	*g ʷk ʷ(ḳ ʷ)* *q ʷ* *(b ʷ f ʷ)* *(m ʷ p ʷ)*
Prepalatals and affricates	*ǧ š* *y*	*š* *y*	*ǧ ɛ* *ɛ̣* *š ʒ* *ñ y*	*ǧ ɛ* *ɛ̣* *š ʒ* *ñ y*	*ǧ ɛ* *ɛ̣* *š ʒ* *ñ y*	*ǧ ɛ* *ɛ̣* *š* *ñ y*	*ǧ ɛ* *ɛ̣* *š ʒ* *ñ y*	*ǧ ɛ* *ɛ̣* *š ʒ* *ñ y* *(q' k' q')*
Laryngeales	*h ḥ* *ʾ ʿ*	*h ḥ* *ʾ ʿ*	*h ḥ* *ʾ ʿ*	*h ḥ* *ʾ ʿ*	*h (ʾ)*	*h ˌ /h/* *ʾ*	*h (ḥ)* */ʾ/*	*(h) (ʾ)*

but *əḫḫelläba* "be milked". As simple (non-geminated) sound, *ḫ* occurs only in Arabic loanwords.

GURAGE. The liquid *l* is extremely rare in Chaha, Eža, Ennemor, Endegeň, and Gyeto. It is replaced in these languages by *n*, *r*.

The sounds *ǩ*, *ǩ*^{w} exist in Western Gurage, except in Gogot; they are absent in Aymellel and in Eastern Gurage.

The rounded labials *b*^{w}, *f*^{w}, *m*^{w}, *p*^{w} exist only in Western Gurage, and in Gafat.

The palatal series *g'*, *k'*, *q'* and *x* is represented mainly in Western Gurage. There are very few examples in Eastern Gurage and in Aymellel.

The laryngeal *h* exists mainly in Eastern Gurage and in Masqan, Gogot, and Endegeň. There are very few examples in the other languages.

The glottal stop ' as representative of an original ' and ' occurs only in Ennemor, Endegeň and Gyeto.[7] As representative of the glottalized *q* it exists in all the dialects except, as it seems, in Chaha, Masqan, and Selti. As representative of the glottalized *ḷ* and *č̣* it also exists in Ennemor and Endegeň.

CORRESPONDENCES. In the correspondence between the Arabic and the Ethiopic phonemes one will notice that quite often several Ethiopic phonemes correspond to one Arabic phoneme. There are various reasons for this situation. The Arabic loanwords could have come into one or several Ethiopic languages from various Arabic dialects, in which case the phonemes would occasionally be different according to the Arabic dialect. Also, the Arabic loanword could have come into one Semitic Ethiopic language through the intermediary of another Semitic Ethiopic language in which the Arabic phonemes are treated differently. Furthermore, the Cushitic languages of the north and south borrowed many words from Arabic, and various Arabic loanwords could have penetrated into Semitic Ethiopic through the intermediary of Cushitic that shows various treatments of the Arabic phonemes. Finally, certain Ethiopic languages, such as Gurage, have different dialects in which the phonemes are treated differently.

For a regular correspondence only two examples are given as

[7] For more details, see *JAOS* 71 (1951), 214

a rule. In a correspondence that is not regular, all the examples are given. The rare occurrence of a correspondence is indicated by the symbols [ ] or ( ). The symbol [ ] indicates a single example for the correspondence; the symbol ( ) is an indication of a rare correspondence, but more than one. The symbol / / indicates that the Ethiopic sound exists only in Arabic loanwords.

Only the consonant correspondences will be considered.

LABIALS

ARABIC *b* = ETHIOPIC *b*.

GEEZ: *zäbib* 'raisin', Ar. *zabīb*; *rəkubat* 'she-camels', Ar. *rakūba*.

TIGRE: *bə'er* 'young camel', Ar. *ba'īr*; *bab* 'dorr', Ar. *bāb*.

TIGRINYA: *bab* 'door', Ar. *bab*; *bäddälä* 'change', Ar. *baddala*.

AMHARIC: *bir* 'cistern', Ar. *bi'r*; *mäšräb* 'water reservoir', Ar. *mašrab*.

HARARI: *bäḵil* 'avaricious', Ar. *baḫīl*; *säbri* 'patience', Ar. *ṣabr*.

ARGOBBA: *räğäb* 'name of month', Ar. *raǧab*; *märḥaba ala* 'agree', Ar. *marḥaba* 'welcome.'

GURAGE: *kəzəb* (C. En. M.) 'lie', Ar. *kaḏaba* 'to lie'; *arb* (E.S.W.Ms.) 'spear, battle', Ar. *ḥarb*.[8]

ARABIC *f* = ETHIOPIC *f*.

GEEZ: *fälək* 'celestial sphere', Ar. *falak*; *ṭənf* 'prominence', Ar. *ṭinf*.

TIGRE: *fäna* 'come to an end', Ar. *fanā*; *fanus* 'lamp', Ar. *fānūs*.

TIGRINYA: *fal* 'omen', Ar. *fāl*; *fanus* 'lamp', Ar. *fānūs*.

AMHARIC: *fəğəl* 'radish', Ar. *fuǧal*; *wäsif* 'adult girl', Ar. *waṣīf*.

HARARI: *fäğri* 'morning', Ar. *faǧr*; *näfa'a* 'serve', Ar. *nafa'a*.

ARGOBBA: *afiya* 'health', Ar. *'āfiya*; *säfär* 'name of month', Ar. *ṣafar*.

GURAGE: *afiya* (Ms. S.) 'health', Ar. *'āfiya*; *afäḷḷärä* (S.) 'break the fast', *afeḷärä* (W.), *afeḷḷärä* (Ms.), Ar. *faṭara*.

ARABIC *m* = ETHIOPIC *m*.

GEEZ: *mi'a* 'myrrh', Ar. *mī'a*; *ḥəməl* 'load', Ar. *ḥiml*.

TIGRE: *mäkruh* 'cursed', Ar. *makrūh*; *məsmər* 'nail', Ar. *mismar*.

TIGRINYA: *'adim* 'red hide', Ar. *'adīm*; *mädfə'* 'canon', Ar. *midfa'*.

[8] By internal change *b* can become *w*: *arəw* (Ed.) 'spear' Ar. *ḥarb*.

Amharic: *mäwǧ* 'wave', Ar. *maụǧ*; *märäq* 'soup', Ar. *maraq*.
Harrri: *qima* 'cost', Ar. *qīma*; *murtad* 'renegade', Ar. *murtādd*.
Argobba: *ǧum'a* 'Friday', Ar. *ǧum'a*; *ǧism* 'body', Ar. *ǧism*.
Gurage: *amd* (C. En.) 'thanks', Ar. *ḥamd*; *mäqäs* (Gur.) 'scissors', Ar. *miqaṣṣ*.[9]

ARABIC *w* = ETHIOPIC *w*.

Geez: *wäba* 'malaria', Ar. *wabā'*; *wäqəf* 'bracelet', Ar. *waqf*.
Tigre: *'awäna* 'help', Ar. *'āwana*; *wäǧəh* 'face', Ar. *waǧh*.
Tigrinya: *'əwan* 'time', Ar. *'awān*; *waǧhi* 'good aspect', Ar. *waǧh*.
Amharic: *woba* 'malaria', Ar. *wabā'*; *wäsif* 'adult girl', Ar. *waṣīf*.
Harari: *wäsiyya* 'last will', Ar. *waṣiyya*; *suwak* 'toothbrush', Ar. *suwāk*.
Argobba: *ḫuwol* 'long', Ar. *ḫawīl*; *wäsiyya* 'last will', Ar. *waṣiyya*.
Gurage: *duwa* (S.W.) 'prayer', Ar. *da'wā*; *suwak* (S.W.) 'toothbrush', Ar. *suwāk*.

DENTALS

ARABIC *d* = ETHIOPIC *d*.

Geez: *'ud* 'aloes', Ar. *'ūd*; *dor, där* 'pearl', Ar. *durr*.
Tigre: *'adäb* 'good behavior', Ar. *'adab*; *'id* 'memorial feast', Ar. *'īd* 'festivity'.
Tigrinya: *dabära* 'live long', Ar. *dābara*; *'adäb* 'good manners', Ar. *'adab*.
Amharic: *mändil* 'handkerchief', Ar. *mandīl*; *gəld* 'hide', Ar. *ǧalad*.
Harari: *da'im* 'always', Ar. *dā'im*; *zahid* 'monk', Ar. *zāhid*.
Argobba: *din* 'belief', Ar. *dīn*; *dunya* 'money', Ar. *'addunyā* 'possessions of this world'.
Gurage: *däwi* (S.W.) 'medicament', Ar. *dawā'*; *ibad* (S.W.) 'prayer', Ar. *'ibād*.

ARABIC *ṭ* = ETHIOPIC *ṭ*.

Geez: *zəfṭ* 'pitch', Ar. *zift*; *zäyṭon* 'olive tree', Ar. *zayṭūn*.
Tigre: *ṭaǧər* 'rich', Ar. *ṭāǧir* 'merchant, rich'; *ṭämma* 'be complete', Ar. *ṭamma*.

[9] By internal change *m* can become *w*: *äwan* (Ed.) 'donkey', Ar. *ḥimār*; *näkkäwä* (Ed.) 'ride', but *läggʷämä* (M. Ms.), *näkʷämä* (E. Gt.), denominative of *ləgʷam* 'bridle', Ar. *liǧām*.

TIGRINYA: *lä'abä* 'be tired', Ar. *la'iba*; *lämämä* 'be completed', Ar. *lamma*.

AMHARIC: *lərəngo* 'cedrat', Ar. *'utrunǧ*; *fälläšä* 'search', Ar. *fallaša*.

HARARI: *lin* 'fig tree', Ar. *līn*; *yälim* 'orphan', Ar. *yatīm*.

ARGOBBA: *wäql* 'time', Ar. *waqt*.

GURAGE: *källäbä* (S.W.) 'write', Ar. *kataba*; *sələr* (C.E.Gt.M.) 'white veil for women', Ar. *sitr*.

ARABIC *ṭ* = ETHIOPIC *ṭ*.

GEEZ: *ṭalisan* 'cloth of priest', Ar. *ṭaylisān*; *baluṭ* 'oak tree', Ar. *ballūṭ*.

TIGRE: *ṭa'a* 'be obedient', Ar. *ṭā'a*; *ṭälläqä* 'be free', Ar. *ṭalaqa*.

TIGRINYA: *qəṭran* 'tar', Ar. *qiṭrān*; *fuṭa* 'piece of cloth', Ar. *fūṭa*.

AMHARIC: *säliṭ* 'sesame', Ar. *salīṭ*; *ṭäbay* 'nature', Ar. *ṭabā'ī'*.

HARARI: *ṭäläqa* 'divorce', Ar. *ṭalaqa*; *šärṭi* 'bet', Ar. *šarṭ* 'condition'.

ARGOBBA: *ṭuwol* 'long', Ar. *ṭawīl*; *ṭäbib* 'magician', Ar. *ṭabīb*.

GURAGE: *näfṭ* (Gur.) 'rifle', Ar. *naft*; *šärmuṭa* 'prostitute', Ar. *šarmūṭa*; *afäṭṭärä* (S.) 'break the fast', *afeṭärä* (W.), *afeṭṭärä* (Ms.), Ar. *faṭara*.

ARABIC *ḍ* = Geez *d*, [*z*];[10] Tigre *d*, [*ṣ*]; Tigrinya *d*; Amharic *d*, *ḍ*, [*z*]; Harari *d*; Argobba *d*; Gurage, no example.

GEEZ: *qade* 'judge', Ar. *qāḍī*; *fədəl* 'excellent', Ar. *fāḍil*; *zäräbä* 'hit', Ar. *ḍaraba*.

TIGRE: *'arid* 'large', Ar. *'arīḍ*; *da'a* 'be lost', Ar. *ḍā'a*; *də'if* 'weak', Ar. *ḍa'īf*; *təqäṣṣäba* 'be in a rage', probably from Ar. *ġaḍiba* 'be angry'.

TIGRINYA: *rädäyä* 'be obedient', Ar. *raḍiya*.

AMHARIC: *qäbd* 'deposit', Ar. *qabḍ*; *ardəya* 'storage dues', Egyptian-Arabic *'arḍiya*; *rämäḍ* 'hot sand', Ar. *'ar-ramḍā'*; *mäqbuḍ* 'bribery', Ar. *maqbūḍ*; *zäbḍiya* 'police', from Egyptian-Arabic *dabaṭ*, *zabaṭ* 'arrest', classical Arabic *ḍabaṭa*.

HARARI: *dämin* 'guarantor', Ar. *ḍamīn*; *bida'a* 'merchandise', Ar. *biḍā'a*; *da'if* 'weak', Ar. *ḍa'īf*.

ARGOBBA: *ramadan* 'name of month', Ar. *ramaḍān*.

In original roots, Arabic *ḍ* corresponds to Geez *ḍ*; Tigre, Tigrinya *ṣ*; South Ethiopic *ḍ*. Example: G. *ḍämädä* "put the yoke" (Ar. *ḍamada*), Te. *ṣämda*, Tna. *ṣämädä*; Amh. *ḍämmädä*.

[10] For the phoneme *ḍ*, see S. Moscati, *Il sistema consonantico delle lingue semitiche*, Rome, 1954, p. 30-32.

In the Cushitic languages, the classical Arabic phoneme *ḍ* in Arabic loanwords is taken over as *d*.[11] Thus, Somali: *dúrub* 'hit', Ar. *ḍaraba*; *diyāfad* 'hospitality', Ar. *ḍiyāfa*; *qādi* 'judge', Ar. *qāḍī*. Bilin: *addāha* 'midday', Ar. *'aḍḍuḥā*. Bedja: *ard* 'land', Ar. *'arḍ*; *demin* 'guarantee', Ar. *ḍamīn*; *wáda* 'religious washing', Ar. *waḍū'*.

As for classical Arabic *ḍ* in the spoken dialects, it appears as emphatic *ḏ* in the Beduin dialects and in Tunisia, as *ḍ* elsewhere, and as *ḷ* in Morocco. Among the Beduins of East Tunisia and in other parts of North Africa and in Malta it appears as *d*. In Persia-Kurdistan, Moṣul, Mardin, and in Turkey Arabic it appears as *z* or *ẓ*; as *z* in Egyptian Arabic, in certain Turkish loanwords of Arabic origin.[12]

ARABIC *s* = Geez *s*; Tigre *s*, (*š*); Tigrinya *s*, (*š*); Amharic *s*, [š]; Harari *s*; Argobba *s*; Gurage *s*.

GEEZ: *sänam* 'hunch', Ar. *sanām*; *bärsim* 'kind of tree', Ar. *birsīm*; *säräḷan*, *šäräḷan* 'cancer', Ar. *saraṭān*; *šämäz* 'bread of wheat', Ar. *samīd*, *samīḏ*.

TIGRE: *sä'a* 'run', Ar. *sa'ā*; *ǧənəs* 'kind', Ar. *ǧins*; *'aballəš* 'devils', Ar. *'iblīs*; *šäḥara* 'bewitch', Ar. *saḥara*.

TIGRINYA: *säliḷ* 'sesame', Ar. *salīṭ*; *'əsli* "swarm', Ar. *'asal*; *läliš* 'mat', Ar. *lillīs*; *šuq̄*, *suq̄* 'market', Ar. *sūq*.

AMHARIC: *sir* 'strip of leather', Ar. *sayr*; *bäläsan* 'balm', Ar. *balasān*; *suq*, *šuq* 'market', Ar. *sūq*.

HARARI: *sullam* 'ladder', Ar. *sullam*; *miskin* 'poor', Ar. *miskīn*.

ARGOBBA: *risala* 'letter', Ar. *risāla*; *ǧism* 'body', Ar. *ǧism*.

GURAGE: *säbäb* (Gur.) 'cause', Ar. *sabab*; *suwak* (S.W.) 'toothbrush', Ar. *siwāk*; *sum* (S.W.) 'poison', Ar. *summ*.

In the original roots, Arabic *s* corresponds to Ethiopic *s*. Example : G. *säm'a* 'hear' (Ar. *sami'a*), Te. *säm'a*, Tna. *säm'e*, Amh. *sämma*.

ARABIC *ṣ* = Geez *ṣ*, *s*; Tigre *s*, (*š*), [*ṣ*]; Tigrinya *s*, *š*, (*ṣ*); Amharic *s*, [š]; Harari *s*; Argobba *s*, *s*^{w}; Gurage *s*.

GEEZ: *ṣa'əf* 'unfermented wine', Ar. *ṣa'f*; *ṣor* 'small palm', Ar. *ṣaḏr*; *säñobar* 'pine tree', Ar. *ṣanābir*; *sädäf* 'pearl', Ar. *ṣadaf*.

[11] The sound *ḍ* does not exist in Bilin. In Somali and Bedja there is a cacuminal *d*.

[12] Kampffmeyer, art. "Arabia. Arabic dialects", *Encyclopedia of Islam*; C. Brockelmann, *Grundriss der vergleichenden Grammatik der semitischen Sprachen*, vol. 1, p. 131.

TIGRE: *'assubuḥ* 'morning', Ar. *'aṣṣubḥ*; *'asäl* 'generation', Ar. *'aṣl*; *'asfär* 'yellow', Ar. *'aṣfar*; *ṣäwwära* 'make a picture', Ar. *ṣawwara*; *qämis* 'garment', Ar. *qamīs*; *šəḥan* 'plate', Ar. *ṣaḥan*.

TIGRINYA: *säfärä* 'be hungry', Ar. *ṣafara* 'be empty'; *sändäl* 'odoriferous wood', Ar. *ṣandal*; *qämis*, *qämiš* 'shirt', Ar. *qamīṣ*; *räsas*, *räšaš* 'lead', Ar. *raṣāṣ*; *wäšəf* 'adult', Ar. *waṣīf*; *ḥərṣät* 'envy', Ar. *ḥirṣ*; *ṣämämä* 'become deaf', Ar. *ṣamma*.

AMHARIC: *qämis* 'shirt', Ar. *qamīṣ*; *suf* 'flannel', Ar. *ṣūf*; *sähan*, *sän* 'plate, dish', Ar. *ṣaḥan*; *säräfä* 'change money', Ar. *ṣarafa*.

HARARI: *saḥib* 'friend', Ar. *ṣāḥib*; *näsib* 'lot', Ar. *naṣīb*; *nəsfi* 'half', Ar. *nuṣf*.

ARGOBBA: *säfär* 'name of month', Ar. *ṣafar*; *wäsiyya* 'last will', Ar. *waṣiyya*; *sʷäbiyy* 'baby', Ar. *ṣabiyy*; *sʷädäqa* 'death commemoration', if it is to be compared with Ar. *ṣadaqa* 'alms, charity'.

GURAGE: *säbi* (Ms.) 'baby', Ar. *ṣabiyy*; *asər* (C.) 'four o'clock in the afternoon', Ar. *'aṣr*; *sanä* (Z.) 'make', Ar. *ṣani'a*.

In original roots, Arabic *ṣ* corresponds to Geez, Tigre, Tigrinya *ṣ*, South Ethiopic *ṭ*.[13] Example: G. *ṣärḥa* 'shout' (Ar. *ṣaraḥa*), Te. *ṣärḥa*, Tna. *ṣärḥe*, Amh. *ṭärra*.

In the Cushitic languages, the phoneme *ṣ* in the Arabic loanwords is taken over as *s*, seldom as *š*.[14] Thus, Somali: *sabi* 'servant', Ar. *ṣabiyy*; *sibāq* 'color', Ar. *ṣibāġ*; *sidri* 'breast', Ar. *ṣadr*. Galla: *risassi* 'bullet', Ar. *raṣāṣ*; *subi* 'dawn', Ar. *ṣubḥ*; *sahani* 'dish', Ar. *ṣaḥn*. Bilin: *asfar* 'yellow', Ar. *'aṣfar*; *qamis* 'shirt', Ar. *qamīṣ*; *sadar* 'breast', Ar. *ṣadar*; *resas* 'lead', Ar. *raṣāṣ*; with *š* in Bilin: *šum* 'fast', either from Ar. *ṣaum* or from Geez *ṣom*. Bedja: *nesir* 'help', Ar. *naṣr*; *asfar* 'yellow', Ar. *'aṣfar*; *asir* 'afternoon', Ar. *'aṣr*; *basal* 'onion', Ar. *baṣal*; with *š* in *ašo* 'adversary', Ar. *'āṣiⁿ*.

Classical Arabic *ṣ* appears also as *s* in the Beduin dialects of Eastern Tunisia, in other parts of North Africa, and in Malta.[15]

In the South Ethiopic languages that do not have the glottalized *ṣ*, the correspondence between Arabic *ṣ* and South Ethiopic *s* is less difficult to explain. The situation is different for Geez, Tigre, and Tigrinya, languages that have the glottalized *ṣ*, and yet Arabic

[13] The affricate *ṣ* (*tˢ'*) is used in North Amharic and in Gafat.

[14] The sound *ṣ* does not exist in Galla, Somali, and Bedja. In Bilin it occurs only in Tigre loanwords (L. Reinisch, *Die Bilin-Sprache in Nordost Afrika*, p. 11).

[15] Kampfmeyer, *op. cit.*

ṣ is often represented as *s* in these languages. Various explanations are possible for this phenomenon. First, the North Ethiopic ṣ is a glottalized sound whereas Arabic ṣ is velarized. These sounds are, therefore, phonetically quite different. Another explanation could be sought in the Cushitic languages of Agau (Bilin, Bedja) and in Galla and Somali. Indeed, these languages have no glottalized or velarized sounds at all, and in the case of any loanword with ṣ (velarized or glottalized) the correspondence is *s*, *š*. It is, therefore, possible that the sound *s* as representative of Arabic ṣ is taken over from Cushitic.

ARABIC *z* = ETHIOPIC *z*.

Geez: *zänabir* 'wasp', Ar. *zanābir*; *zən(n)ar* 'belt', Ar. *zunnār*.

Tigre: *näzla* 'go down', Ar. *nazala*; *zänbil* 'basket', Ar. *zanbīl*.

Tigrinya: *zäbib* 'raisin', Ar. *zabīb*; *qäzärä* 'be abundant', Ar. *ǧazura*.

Amharic: *zäbib* 'raisin', Ar. *zabīb*; *zəbad* 'civet cat', Ar. *zabād*.

Harari: *ziyara* 'visit', Ar. *ziyāra*; *zayid* 'excess', Ar. *zā'id*.

Argobba: no example.

Gurage: *ǧeza* (C. En.M.) 'gratitude', Ar. *ǧazā'*; *zənnar* (Gur.) 'cartridge poach', Ar. *zunnār*.[16]

INTERDENTALS

ARABIC *ḏ* = Geez *d*, *z*; Tigre *d*, *z*; Tigrinya *d*, *z*; Amharic *z*, [*d*], [*l*]; Harari *z*; Argobba *z*; Gurage [*d*], [*z*].

Geez: *'udr* 'excuse', Ar. *'uḏr*; *qädäfä*, *qäzäfä* 'row', Ar. *qaḏafa*; *šämäz* 'bread of wheat', Ar. *samīd*, *samīḏ*; *lozan* 'perfume', Ar. *'allāḏūn*.

Tihre: *'ada*, *'aza* 'harm', Ar. *'aḏā*; *'addäba* 'trouble', Ar. *'aḏḏaba*; *zälil* 'powerless', Ar. *ḏalīl*; *läzza* 'be of mild character', Ar. *laḏḏa*.

Tigrinya: *dähab* 'gold', Ar. *ḏahab*; *tälmedan* 'student worker', Ar. *tilmīd*; *lazza* 'flavor, taste', Ar. *laḏḏa*; *qäzäfä* 'row', Ar. *qaḏafa*.

Amharic: *qäzzäfä* 'row', Ar. *qaḏafa*; *läzza* 'pleasant taste', Ar. *laḏḏa*; *zänäb* 'hall', Ar. *ḏanab* 'tail'; *dura* 'meter', Ar. *dirā'*; *näbit* 'wine', Ar. *nabīḏ*.

Harari: *əzni* 'permission', Ar. *'iḏd*; *käzzab* 'liar', Ar. *kaḏḏāb*.

[16] In the various Gurage dialects, *z* is occasionally palatalized into *ž* and alternates with *ǧ*: *anž*, *añž*, *aǧ* (C. En. Gt. Ed. Ms.) 'calf', Ar. *'anz*; *zənǧəbəl* (W.) 'ginger', *ǧanǧəbəl* (S.), *ǧənǧəbəl* (A.), Ar. *zanǧabīl*.

ARGOBBA: *zil qi'da* 'name of month', Ar. *ḏū-l-qa'da*; *kisł* 'lie', for *kiz-ł*, **kzb*, Ar. *kiḏb*.

GURAGE: *edäbä* (W.) 'be in need', *idäbä* (Z.), *aliddäbä* (A.) 'upset', Ar. *'aḏḏaba* 'punish'; *kəzəb* (Z.En.M.) 'lie', *kizb* (S.W.), Ar. *kaḏaba*.

In native roots, Arabic *ḏ* corresponds to Ethiopic *z*. Example: G. *'əzən* 'ear' (Ar. *'uḏn*), Te. *'əzən*, Tna. *'əzni*; G. *zäkärä* 'remember' (Ar. *ḏakara*), Te. *zäkra*, Tna. *zäkkärä* 'remind', Amh. *täzäkkärä* 'remember'.

In the Cushitic languages, the interdental *ḏ* of the Arabic loanwords is taken over as *d*.[17] Thus, Somali: *idin* 'permission', Ar. *'iḏn*; *'adāb* 'pain', Ar. *'aḏāb*; *nidar* 'vow', Ar. *naḏr*. Bilin: *'adäb* 'cause pain', Ar. *'aḏḏaba*; *dahab* 'gold', Ar. *ḏahab*; *qadaf* 'oar', Ar. *qaḏafa*.

In the spoken Arabic dialects, *ḏ* is preserved in Oman, Ḥaḍramaut, Mosul, Mardin, in the Beduin dialects of Neǧd, Syria and Tunisia. In the other dialects it became *d*. In Egyptian Arabic it is also represented occasionally as *z* and *d*.[18]

ARABIC *ɬ*=Geez *ł*, *s*; Tigre *ł*, [s]; Tigrinya [*ł*], *s*; Amharic [*ł*], [*s*]; Harari *ł*, *s*; Argobba [*s*]; Gurage [*ł*].

GEEZ: *tərya* 'pleiades', Ar. *ɬurayya*; *'awłan* 'idols', Ar. *'awɬān*; *säwr*, *täwr* 'bull as sign of Zodiac', Ar. *ɬaur*; *som*, *somät*, *łuma* 'garlic', Ar. *ɬūm*.

TIGRE: *'alafə'* 'three stones of the cooking pot', Ar. *'aɬāfiⁿ*; *lädə'* 'breast', Ar. *ɬady*; *'asbäła* 'bring witness', Ar. *'aɬbała*.

TIGRINYA: *łəmni* 'eighth of thaler', Ar. *ɬumn*; *färsi* 'chyme', Ar. *farɬ*; *sor* 'castrated he-goat', Ar. *ɬaur*.

AMHARIC: *łämun* 'eighth of thaler', Ar. *ɬumn*; *färs* 'contents of stomach', Ar. *farɬ*.

HARARI: *łämanīn* 'eighty', Ar. *ɬamānīn*; *łumma* 'garlic', Ar. *ɬūm*; *baḥsi aša* 'investigate', Ar. *baḥɬ*; *sälāsa* 'Tuesday', Ar. *ɬulāɬā'*; *isnīn* 'Monday', Ar. *'al-'iɬnayni*.

ARGOBBA: *rabi sani* 'name of month', Ar. *rabī' ɬāni*.

GURAGE: *łuma*, *łumma* (Gur.) 'garlic', Ar. *ɬūm*.

In native roots, Arabic *ɬ* corresponds to Ethiopic *s*. Examples: G. *ḥaräsä* 'plough' (Ar. *ḥaraɬa*), Te. *ḥarsa*, Tna. *ḥaräsä*, Amh. *arräsä*; G. *ḥaddäsä* 'renew' (Ar. *ḥadaɬa*), Te. *ḥaddäsa*, Tna. *ḥaddäsä*, Amh. *addäsä*.

[17] The interdental *ḏ* does not exist in Cushitic.

[18] Brockelmann, *op. cit.* p. 120, 131.

In the Cushitic languages, the interdental *ṯ* of the Arabic loanwords is taken over as *s*, *t*.[19] Somali: *tōn* 'garlic', Ar. *ṯūm*; *masal* 'comparison', Ar. *maṯal*; *salāsa* 'Tuesday', Ar. *ṯalāṯā'*. Galla: *tamuni* 'sixteenth of a thaler', Ar. *ṯumn* 'eighth'. Bedja: *talāta* 'Tuesday', Ar. *ṯalāṯā'*; *temin* 'price', Ar. *ṯaman*.

In the Arabic dialects, the classical *ṯ* is represented as *ṯ* or *t* in the same regions in which classical *ḏ* is represented as *ḏ* or *d* (see above). In Egyptian Arabic, classical Arabic is borrowed as *s*, occasionally as *t*.[20]

ARABIC *ẓ* = Geez [*ẓ*];[21] Tigre *z*, *d*; Tigrinya [*ẓ*]; Amharic [*d*]; Harari *z*; Gurage [*z*].

GEEZ: *ẓälämä* 'do injustice', Ar. *ẓalama*.

TIGRE: *zälläma*, *dälläma* 'do injustice', Ar. *ẓalama*; *nazära* 'inspect', Ar. *naẓara*; *addəhər* 'noon', Ar. *'aẓẓuhr*; *nädfa* 'be clean', Ar. *naẓufa*.

TIGRINYA: *ẓälämä* 'break one's word', Ar. *ẓalama*.

AMHARIC: *mafuda* 'purse', if it is to be compared with Ar. *maḥfūẓa* 'guarded'.

HARARI: *zeläma* 'do injustice', Ar. *ẓalama*; *nazif* 'clean', Ar. *naẓīf*.

In native roots, Arabic *ẓ* corresponds to Geez, Tigre, Tigrinya *ṣ*, South Ethiopic *ḍ*. Examples: G. *ṣäm'a* 'be thirsty' (Ar. *ẓami'a*), Te. *ṣäm'a*, Tna. *ṣäm'e*, Amh, *ḍämma*; G. *'aṣəm* (also *'aḍəm*) 'bone' (Ar. *'aẓm*), Te. *'aṣəm*, Tna. *'aṣmi*, Amh. *aḍənt*.

In the Cushitic languages, the sound *ẓ* of the Arabic loanwords corresponds to *d*.[22] Thus, Somali: *nadif* 'clean', Ar. *naẓīf*; *duhur* 'midday', Ar. *ẓuhr*; *mandar* 'mirror', Ar. *miṇẓar*. Bilin: *nadif* 'clean', Ar. *naẓīf*. Bedja: *duluma* 'darkness', Ar. *ẓulma*; *nadif* 'clean', Ar. *naẓīf*; *din* 'think', Ar. *ẓanna*.

In the spoken Arabic dialects, classical *ẓ* is represented by the same sound as classical *ḏ* (see above). In Persia and in Turkey Arabic it is represented as *z*. In Egyptian Arabic, classical *ẓ* is borrowed as *z*.[23]

[19] The interdental *t* does not exist in Cushitic.

[20] Brockelmann, *op. cit.* 120, 131.

[21] On Arabic *ẓ*, see S. Moscati, *Il sistema consonantico delle lingue semitiche*, p. 29-30, and *passim*.

[22] The Cushitic languages have no *ẓ*. Somali and Bedja have no *z*.

[23] Kampfmeyer, *op. cit.*; Brockelmann, *op. cit.* 124, 131.

LIQUIDS

ARABIC *n* = ETHIOPIC *n*.

GEEZ: *nora* 'chalk', Ar. *nūra*; *mänarät* 'candle-stick', Ar. *manāra*.
TIGRE: *näf'a* 'be useful', Ar. *nafa'a*; *nada* 'announce', Ar. *nadā*.
TIGRINYA: *mändil* 'kerchief', Ar. *mandīl*; *'addənya* 'multitude', Ar. *'addunyā*.
AMHARIC: *ɬanika* 'tin', Ar. *ɬanika*; *näbiɬ* 'wine', Ar. *nabīḍ*.
HARARI: *nəsfi* 'half', Ar. *nuṣf*; *qärni* 'persons of the same age', Ar. *qarni*.
ARGOBBA: *mänam* 'dream', Ar. *manām*; *din* 'belief', Ar. *dīn*.
GURAGE: *ǧin* (W.) 'spirit', Ar. *ǧinn*; *ǧäbän* (Gt.) 'jug', Ar. *ǧaban*.[24]

ARABIC *l* = ETHIOPIC *l*.

GEEZ; *ləfaf* 'envelop', Ar. *lifāfa*; *qälläbä* 'turn over', Ar. *qalaba*.
TIGRE: *'awwäl* 'formerly', Ar. *'awwal*; *ləǧam* 'bridle', Ar. *liǧām*.
TIGRINYA: *'amil* 'client', Ar. *'āmil*; *'addälä* 'distribute', Ar. *'addala*.
AMHARIC: *fəǧəl* 'radish', Ar. *fuǧal*; *säliṭ* 'sesame', Ar. *salīṭ*.
HARARI: *lākin* 'but', Ar. *lākin*; *šuqli* 'work', Ar. *šuǧl*.
ARGOBBA: *risala* 'letter', Ar. *risāla*; *ǧism* 'body', Ar. *ǧism*.
GURAGE: *alla* (S. W.) 'God', Ar. *'allah*; *ləbab* (M.Ms.A.) 'head-stall', Ar. *libāb*.[25]

ARABIC *r* = ETHIOPIC *r*.

GEEZ: *qərd* 'monkey', Ar. *qird*; *märkäb* 'boat', Ar. *markab*.
TIGRE: *'əmər* 'life', Ar. *'umr*; *räda* 'consent, agree', Ar. *raḍā*.
TIGRINYA: *'əmri* 'life', Ar. *'umr*; *bärädä* 'polish', Ar. *barada*.
AMHARIC: *šorba* 'soup', Ar. *šurba*; *aräqi* 'liquor', Ar. *'araqi*.
HARARI: *räṭṭi* 'pound', Ar. *raṭl*; *sirri* 'secret', Ar. *sirr*.
ARGOBBA: *harb* 'spear', Ar. *ḥarb* 'battle'; *märḥaba ala* 'agree', Ar. *marḥaba* 'all right'.
GURAGE: *aḵera* (C.En.) 'future life', Ar. *'al-'āḫira*; *ker* (C.En.A.) 'well, good', Ar. *ḫayr*.[26]

[24] By an internal phonetic process also *r, ñ*: *mezänä* (S.) 'weigh, count', *mezärä* (C. En. Gt.), *mezzärä* (S.), denominative of *mizan* 'balance', Ar. *mizān*; *arbaň* (M. A.) 'hare', *arbaňňa* (S. W.), *arwaňňā* (Go.), if it is borrowed from Ar. *'arnab*.

[25] By an internal process also *n, r*; *baliq* (Ed. Ms. S. W.) 'old man', *bariq* (C. E. En. Gt.), Ar. *bāliġ*; *lägg'ämä* (M. Ms.) 'ride', *näk'ämä* (C. Gt.), *nägg'ämä* (E.), from *lägᵛam* 'bridle', Ar. *liǧām*.

[26] By an internal process also *n*: *qärä* (S. W.) 'study, read the Koran', *qäna* (En.), from Ar. *qara'a*.

PREPALATALS AND AFFRICATES

ARABIC *š* = Geez *š* also *s*; the other languages *š*.

GEEZ: *šahər* 'new moon', Ar. *šahr*; *šärg* 'rivulet', Ar. *šarǧ*; *säbdə'al* 'viper', Ar. *šibdi'*; *gäwsan* 'cuirass', Ar. *ǧaušan*; *sänʿäräǧ*, *sänʿäräž* 'chess', Ar. *šiʿranǧ*.

TIGRE: *fəraš* 'carpet', Ar. *firāš*; *šäka* 'complain', Ar. *šakā*.

TIGRINYA: *šaš* 'muslin', Ar. *šāš*; *fərraš* 'matrass', Ar. *firāš*.

AMHARIC: *šälala* 'paralyzed', Ar. *šalla* 'be paralyzed'; *fäläšä* 'verify, search', Ar. *falaša*.

HARARI: *šarīk* 'participant', Ar. *šarīk*; *minšār* 'saw', Ar. *minšār*.

ARGOBBA: *šäwwal* 'name of month', Ar. *šawwāl*; *abšər* 'courage!', Ar. *'abšir* 'bring good news!'.

GURAGE: *šaš* (Gur.) 'muslin', Ar. *šāš*; *šahid*, *šayəd* 'witness', Ar. *šāhid*.

ARABIC *ǧ* = Geez *g*, g^w, late *ǧ*, *ž*; Tigre *ǧ*, *g*; Tigrinya *ǧ*, *g*, g^w; Amharic *ǧ*, *g*, g^w; Harari *ǧ*, (*g*); Argobba *ǧ*; Gurage *ǧ*, *g*, g^w, /*g'*/ in Western Gurage.

GEEZ: *gäziräl* 'island', Ar. *ǧazīra*; *mäsgid* 'mosque', Ar. *masǧid*; g^w*əbnäl*, *gəbnäl* 'cheese', Ar. *ǧubna*; g^w*ərn* 'threshing field', Ar. *ǧurn*; *ǧäyb* 'pocket', Ar. *ǧayb*; *sänʿäräg*, *sänʿäräž* 'chess', Ar. *šiʿranǧ*; *zīž* 'astronomical tables', Ar. *zīǧ*; *'əlʿažra* 'the Hidjra', Ar. *'al-hiǧra*; *žädəy*, *gädəy* 'capricorn as sign of Zodiac', Ar. *ǧady*.

TIGRE: *lə 'aǧǧäba* 'wonder', Ar. *taʿaǧǧaba*; *ǧära* 'happen', Ar. *ǧarā*; *gäd* 'luck', Ar. *ǧadd*; *näga* 'be in good health', Ar. *naǧā*.

TIGRINYA: *ǧimla* 'together', Ar. *ǧumla*; *ǧärräbä* 'try', Ar. *ǧarraba*; *gämri* 'in mass', Ar. *ǧamār*; *gäddi* 'luck', Ar. *ǧadd*; *lərəng*^{w}*i* 'cedrat', Ar. *'utrunǧ*; *ləg*^{w}*am* 'bridle', Ar. *liǧām*.

AMHARIC: *fəǧəl* 'radish', Ar. *fuǧal*; *ǧeb* 'pocket', Ar. *ǧayb*; *gəld* 'hide', Ar. *ǧalad*; *gəbli* 'mountainous', Ar. *ǧabaliyy*; *aggämä* 'to cup', Ar. *ḥaǧama*; *ləg*^{w}*am* 'bridle', Ar. *liǧām*.

HARARI: *ǧinsi* 'kind', Ar. *ǧins*; *liǧāra* 'trade', Ar. *liǧāra*; *ḥāǧa* 'need', Ar. *ḥāǧa*; *gädid* 'fresh', Ar. *ǧadīd*; *ḥangur* 'throat', Ar. *ḥanǧūr*.

ARGOBBA: *ǧism* 'body', Ar. *ǧism*; *ǧum'a* 'Friday', Ar. *ǧumʿa*.

GURAGE: *ǧeza* (C.En.M.) 'gratitude', Ar. *ǧazā'*; *ǧənez* (En.) 'corpse', *ǧinayz* (S.), Ar. *ǧanā'iz*; *agäbärä* (S.) 'tame', Ar. *ǧabara* 'force'; *agämä*, *aggämä* (Gur.) 'to cup', Ar. *ḥaǧama*; *ləg*^{w}*am*, *lugam*

(Gur.) 'bridle', Ar. *liǧām*; *lägg'ämä* (M.Ms.) 'ride', *nägg'ämä* (E.), denominative of *ləgʷam* 'bridle', Ar. *liǧām*.

In native roots, Arabic *ǧ* corresponds to Ethiopic *g*. Example: G. *gämäl* 'camel' (Ar. *ǧamal*), other Ethiopic languages *gämäl*.

In the Cushitic languages, the sound *ǧ* of the Arabic loanwords is taken over as *ǧ*, *g*. Thus, Galla : *ǧimata* 'Friday', Ar. *ǧum'a*; *ǧelba* 'boat', Ar. *ǧalba*; *masgid* 'mosque', Ar. *masǧid*; *naga* 'peace', Ar. *naǧā* 'be safe'. Bilin: *ǧalba* 'small boat', Ar. *ǧalba*; *hoǧeb* 'eyebrow', Ar. *ḥāǧib*; *ǧänäbäl* and *gänäbäl* 'sin', Ar. *ǧanaba*; *genadäl* 'corpse', Ar. *ǧanāza*. Bedja: *gibne*, in the South *ǧibne* 'cheese', Ar. *ǧubna*; *gum'a* and *ǧum'a* 'Friday', Ar. *ǧum'a*; *gamus* and *ǧamus* 'buffalo', Ar. *ǧāmūs*.

In the spoken Arabic dialects, classical *ǧ* is represented as *ǧ* in Iraq, partly in Central Arabia, Mecca, Aleppo, and northern Algeria; as *g* in Lower Egypt (particulary in Cairo), and partly in South Arabia; as *ž* in Jerusalem, in the towns of the Syrian coast, in certain districts of Lebanon, Damascus, Mesopotamia, Tunisia, Tripolitania, Morocco, and South Algeria.[27]

ARABIC *y* = ETHIOPIC *y*.

GEEZ: *ləry a* 'pleiades', Ar. *ṯurayya*; *qäyṯun* 'small nook in the house', Ar. *qayṯūn*.

TIGRE: *qäyyäda* 'fetter', Ar. *qayyada*; *yätim* 'orphan', Ar. *yatīm*.

TIGRINYA: *qäyyäsä* 'measure', Ar. *qayyasa*; *qäyyädä* 'tie', Ar. *qayyada*.

AMHARIC: *ṯäyar* 'dromedary', Ar. *ṯayyār*; *säyara* 'caravan', Ar. *sayyāra*.

HARARI: *yätim* 'orphan', Ar. *yatīm*; *ḫiyāna* 'treachery', Ar. *ḫiyāna*.

ARGOBBA: *wäsiyya* 'last will', Ar. *waṣiyya*; *afiya* 'health', Ar. *'āfiya*.

GURAGE: *afiya* (Ms.S.) 'health', *ayfa* (C.En.), Ar. *'āfiya*; *aynät* (M.Ms.A.S.W.) 'kind', Ar. *'ayna*.

VELARS

ARABIC *k* = Geez *k*, *kʷ*; Tigre *k*; Tigrinya *k*, *kʷ*, *ḵ* (intervocalic); Amharic *k*, *kʷ*; Harari *k*, (*ḵ*); Argobba *k*; Gurage *k*, *kʷ*.

GEEZ: *käbärit* 'sulphur', Ar. *kibrīt*; *wäkäbä* 'do something assiduously', Ar. *wakaba*; *kʷəḥəl* 'antimony', Ar. *kuḥl*.

[27] W. Marçais, art. 'Djim', *Encyclopedia of Islam*; Brockelmann, *op. cit.* 122, 123.

TIGRE: *ḥaka* 'tell', Ar. *ḥakā*; *kämal* 'perfection', Ar. *kamāl*.

TIGRINYA: *kərbil* 'match', Ar. *kibrīt*; *käsäbä* 'gain in business', Ar. *kasaba*; *kʷəḥli* 'antimony', Ar. *kuḥl*; *wäḵil* 'agent', Ar. *wakīl*; *rəḵab* 'stirrup', Ar. *rikāb*.

AMHARIC: *hakim* 'physician', Ar. *ḥakīm*; *kämun* 'cumin', Ar. *kammūn*; *kʷəl* 'antimony', Ar. *kuḥl*; *dənkʷan* 'tent', Ar. *dukkān*.

HARARI: *kursi* 'chair', Ar. *kursī*; *kätäbä* 'write', Ar. *kataba*; *käsbi*, *ḵäsbi* 'profit', Ar. *kasb*; *šubbak*, *šubbaḵ* 'window', Ar. *šubbāk*.[28]

ARGOBBA: *kəbər* 'rich', Ar. *kabīr*; *kisl* 'lie', for *kiz-l*, **kzb*, Ar. *kiḏb*.

GURAGE: *kəzəb* (C.En.M.) 'lie', Ar. *kiḏb*; *kätäbä* (S.W.) 'write', Ar. *kataba*; *kor*, *kʷär* (Gur.) 'saddle', Ar. *kūr*; *kofilc*, *kʷäfila* (C.En.) 'plaited turban of Moslems', Ar. (Syrian) *kūfiya*.

ARABIC *q* = Geez *q*, *qʷ*; Tigre *q*, (*g*); Tigrinya *q*, *qʷ*, *q̄* (intervocalic); Amharic *q*, *qʷ*; Harari *q*, [*k*]; Argobba *q*; Gurage *q*.

GEEZ: *naqät* 'she-camel', Ar. *nāqa*; *qəṭran* 'liquid pitch', Ar. *qiṭrān*; *näqʷəṭ* 'dot', Ar. *nuqṭa*.

TIGRE: *'aqəl* 'wise', Ar. *'āqil*; *qadər* 'powerful', Ar. *qādir*; *gäbilät* 'tribe', Ar. *qabīla*; *gäsma* 'divide', Ar. *qasama*.

TIGRINYA: *qəflät* 'caravan', Ar. *qāfila*, *qälbi* 'reason', Ar. *qalb* 'heart'; *həqqʷa* 'snuff box', Ar. *ḥuqqa*; *šəq̄ʷi* 'market', Ar. *sūq*; *'aq̄li* 'patience', Ar. *'aql*; *mäzräq̄* 'javelin', Ar. *mizrāq*.

AMHARIC: *aräqi* 'liquor', Ar. *'araqi*; *qumru* 'kind of pigeon', Ar. *qumrī*; *həqqʷa* 'snuff box', Ar. *ḥuqqa*; *näqʷəṭ* 'four dots at the end of the sentence', Ar. *nuqṭa*.

HARARI: *qiwwa* 'strength', Ar. *quwwa*; *qälbi* 'heart', Ar. *qalb*; *märäḵ* 'soup', Ar. *maraq*.

ARGOBBA: *wäqt* 'time', Ar. *waqt*; *qarra* 'read', Ar. *qara'a*.

GURAGE: *aq* (W.Go.) 'truth', *anq* (C.En.Ed.), *haq* (S.), Ar. *ḥaqq*; *qärä* (S.W.) 'study, read the Koran', *qärra* (Ms.), Ar. *qara'a*.

It is interesting to note that there are no examples of Arabic *q* appearing in the loanwords as '. And yet, classical *q* is represented as ' in certain regions of Egypt (Cairo, Fayyum), in the majority of cities of Syria, in North Morocco, and in Maltese. It appears voiced in Ḥaḍramout, Oman, and in some Beduin dialects; as *k* among the Beduins of Eastern Tunisia, in certain parts of North Africa, in Malta, and in the surroundings of Jerusalem.[29]

[28] The sound *ḵ* of Harari is often an allophone of *k*. Somali, too, occasionally renders Arabic *k* by *ḵ*; thus, *saḵran*, 'drunk', Ar. *sakrān*.

ARABIC *ḫ* = Geez *k*, *kʷ*, *ḫ*, late *h*; Tigre *k*, (*ḫ*); Tigrinya *k*, *kʷ*, *ḵ* (intervocalic or after vowel), (*ḫ*); Amharic *k*, *kʷ*, (*h*); Harari *ḵ*, *k*, (*ḫ*); Argobba *k*, [*ḵ*]; Gurage *k*, *ḵ*, (*h*), *zero*.

GEEZ: *käbäzä* 'bake', Ar. *ḫubz* 'bread';[30] *kämər* 'wine', Ar. *ḫamr*; *rəkʷam* 'marble', Ar. *ruḫām*; *həwur* 'weak', Ar. *ḫawwār*; *ḫäräṣil* 'purse', Ar. *ḫurṣ*; *säliḫoł*, *sälik* 'cassia plant', Ar. *salīḫa*.

TIGRE: *'ākər* 'end', Ar. *'āḫir*; *käbär* 'notice', Ar. *ḫabar*; *kälqa* 'create', Ar. *ḫalaqa*; *ḥarba* 'fall in ruins', Ar. *ḫaraba*; *ḥaṭra* 'be courageous', Ar. *ḫaṭara*; *ḥarəb* 'waterskin', Ar. *ḫurba*, *ḫurba*.

TIGRINYA: *käbär* 'notice', Ar. *ḫabar*; *ker* 'good', Ar. *ḫayr*; *kʷərği* 'saddle bag', Ar. *ḫurǧ*; *räḵäsä* 'be cheap', Ar. *raḫaṣa*; *kuḵ* 'peach', Ar. *ḫūḫ*; *ḥabbärä* 'make know', Ar. *ḫabbara*; *tä-zaḥarä* 'boast', Ar. *zaḫara*.

AMHARIC: *käl* 'vinegar', Ar. *ḫall*; *mäkädda* 'cushion, pillow', Ar. *miḫadda*; *källäsä* 'sell everything one has', Ar. *ḫalaṣa* 'finish'; *kʷäk*, *kok* 'peach', Ar. *ḫūḫ*; *mäkräǧ*, *mähräǧ* 'kettle', Ar. *ḫurǧa*; *ǧuh*, *žuh* 'red cloth', Ar. *ǧūḫ*.

HARARI: *ḵilāf* 'opposition', Ar. *ḫilāf*; *ḵäläqa* 'create', Ar. *ḫalaqa*; *kuffi*, *ḵuffi* 'shoe', Ar. *ḫuff*; *ḵäsmi*, *käsmäññа* 'adversary', Ar. *ḫaṣm*; *ḵäräzāni*, *käräzāni* 'stick of bamboo', Ar. *ḫayzurān*; *al-ḥebära* 'ask', Ar. *'istaḫbara*; *šäḥał aša* 'urinate', Ar. *šaḫḫa* 'urinate', verbal noun *šaḫḫał*.

ARGOBBA: *kämis* 'Thursday', Ar. *ḫamīs*; *räkis* 'cheap', Ar. *raḫīṣ*; *ḵälläqa* 'create', Ar. *ḫalaqa*.

GURAGE: *ker* (C.En.A.) 'good, well', Ar. *ḫayr*; *käms* (En.Gt. Ed.S.W.) 'Thursday', Ar. *ḫamīs*; *käsärä* (C.En.Gt.S.W.) 'lose in business', Ar. *ḫasara*; *aḵera* (C.En.) 'future life', Ar. *'al-'āḫira*; *ḵäläqä* (W.) 'create', *haläqä* (Ulbarag), *halläqä* (Ms.), *aläqä* (S.), Ar. *ḫalaqa*; *mähräǧ* (W.) 'jug', perhaps from Ar. *ḫurǧa* 'bucket'.

In native roots, Arabic *ḫ* corresponds to Geez *ḫ*; Tigre, Tigrinya *ḫ*; Amharic, Gafat, Gurage *zero*; Harari *ḫ*; Argobba *h*.[31] Examples: G. *ḫaläfä* 'pass' (Ar. *ḫalafa*), Te. *ḫalfa*, Tna. *ḫaläfä*, Amh. *alläfä*, Har. *ḫuluf baya*; G. *ḫanäqä* 'strangle' (Ar. *ḫanaqa*), Te. *ḫanqa*, Tna. *ḫanäqä*, Amh. *annäqä*, Har. *ḫanäqa*, Arg. *hannäqa*.

In the Cushitic languages, the velar spirant *ḫ* is rendered by *k*,

[29] By an internal phonetic process *q* appears as ' in some Gurage dialects. Thus, *qərära* (C. En. Gt. M.) 'early morning', *'ərär'a* (En. Ed.), Ar. *'al-qarralāni*.

[30] This root has a native form in G. *ḫəbəsi* 'bread', for *ḫəbəzi*, with *z* becoming voiceless *s* because of the following voiceless *t*.

[31] For more details about the original *ḫ* in South Ethiopic, see *JAOS* 71 (1951). 213-215.

rarely by *h*. Thus, Bilin: *kadam* 'serve', Ar. *ḫadama*; *ker* 'good', Ar. *ḫayr*; *kaymat* 'tent', Ar. *ḫayma*; *akerā* 'nether world', Ar. *'āḫira*. Bedja; *kedim* 'serve', Ar. *ḫadama*; *keli* 'be free', Ar. *ḫalā*; *keliq* 'create', Ar. *ḫalaqa*; *hamīs* 'Thursday', Ar. *ḫamīs*; *hōš* 'courtyard', Ar. *ḫauš*. Somali: *kibis* 'bread', Ar. *ḫubz*; *keli* 'alone', Ar. *ḫalw*; *kātun* 'seal', Ar. *ḫātim*. Galla: *kamisa* 'Thursday', Ar. *ḫamīs*; *kasara* 'bankruptcy', Ar. *ḫasāra*; *rakasa* 'cheap', Ar. *raḫīs*. Sidamo: *kalqi* 'create', Ar. *ḫalaqa*; *käms* (Alaba) 'Thursday', Ar. *ḫamīs*; *kere* 'well', Ar. *ḫayr*.

It is quite possible that the correspondence of Semitic Ethiopic *k* to *ḫ* of the Arabic loanwords in due to the influence of Cushitic.

As it can be seen from the list of the phonemes, all the Ethiopic languages have the labiovelar series except Tigre and Harari.[32] In most of the cited loanwords the Ethiopic labiovelar is due to the presence of the vowel *u* in the original words even though this vowel no longer appears in the Ethiopic word. One finds, however, a labiovelar in Ethiopic also in examples where there is no vowel *u* in the original form of classical Arabic. In all these cases it is possible either that the borrowings come from an Arabic dialect that had a labial vowel (*u*, *o*) in the root, or that the labiovelar being a regular feature of the Ethiopian phonology appears in borrowings even though there is no labial vowel that would have caused the labialization.

The examples are: Geez *gʷəbnät* 'cheese', Ar. *ǧubna*; *dəgʷəlam* 'bracelet', Ar. *dumluǧ*; *ləgʷam* 'bridle', Ar. *liǧām*, *luǧām*; *kʷəḥəl* 'antimony', Ar. *kuḥl*; *rəkʷam* 'marble', Ar. *ruḫām*; *qʷəsṭ*, *qʷäsṭ*, *qusṭ* 'aromatic plant', Ar. *qusṭ*; *näqʷəṭ* 'dot', Ar. *nuqṭa*; *qʷəṭən* 'silk', Ar. *quṭn*. Without the vowel *u* in the original root: *dəgʷər* 'wood of the plough', Ar. *daǧr*; *ḥaqʷänä* 'churn', Ar. *ḫaqana*; *qʷəsyäl*, *qʷəsäyät* 'cucumber', Ar. *qiṯṯa'*. Tigrinya: *kʷərǧi* 'saddle bag', Ar. *ḫurǧ*; *ləgʷam* 'bridle', Ar. *liǧām*, *luǧām*; *kʷəḥli* 'antimony', Ar. *kuḥl*; *šəqʷ* 'market', Ar. *sūq*; *həqqʷa* 'snuff box', Ar. *ḫuqqa*; *näqʷṭi* 'four dots at the end of the sentence', Ar. *nuqṭa*; *tərəngʷi* 'cedrat', Ar. *'utrunǧ*; *qʷäränä* 'to yoke', probably from **qurn* 'yoke' rather than from *qarana*. Amharic: *ləgʷam* 'bridle', Ar. *liǧām*, *luǧām*; *dənkʷan* 'tent', Ar. *dukkān*; *kʷəl* 'antimony', Ar. *kuḥl*; *rəkʷam*, *rukam* 'marble', Ar. *ruḫām*; *kok*, *kʷäk* 'peach', Ar. *ḫūḫ*; *qʷəran* 'Koran', Ar. *qur'ān*; *həqqʷa* 'snuff box', Ar. *ḫuqqa*; *näqʷəṭ* 'four dots at the end of the sentence', Ar. *nuqṭa*. Gurage:

[32] On the labiovelars, see also E. Ullendorff, *The Semitic languages of Ethiopia. A comparative phonology*, p. 74ff.

ləgʷam, *lugam* 'bridle', Ar. *liǧām*; *kor*, *kʷär* 'saddle', Ar. *kūr*; *kofila*, *kʷäfila* (C.En.) 'plaited turban of the Moslems', Ar. *kūfiya*.

ARABIC *ġ* = Geez *q*; Tigre *q*; Tigrinya *q*, *q̄* (intervocalic); Amharic *q*, *qʷ* [*g*]; Harari *q*, *ḵ*, /*ġ*/, [*g*]; Argobba [*q*]; Gurage *q*, [*g*].

GEEZ: *bälqäm* 'phlegma', Ar. *balġam*.[33]

TIGRE: *qafläl* 'suddenly', Ar. *ġafla*; *šəqəl* 'work', Ar. *šuġl*; *qäla* 'be expensive', Ar. *ġalā*.

TIGRINYA: *qali* 'expensive', Ar. *ġālī*; *qäyyärä* 'change', Ar. *ġayyara*; *šəq̄läl* 'work', Ar. *šuġl*.

AMHARIC: *tä-šaqqälä* 'be occupied', Ar. *šuġila*; *balläqä* 'become adult', Ar. *bālaġa*; *qʷäle*, *qole*, *qule* 'spirit', Ar. *ġūl*; *gərš* 'small coin', from Ar. *ġurš* through *Groschen*.

HARARI: *šuqli* 'work', Ar. *šuġl*;[34] *qali* 'expensive', Ar. *ġālī*; *sibāḵ* 'color', Ar. *ṣibāġ*; *istifrāḵ āša* 'vomit', Ar. *'istifrāġ*, from *'istafraġa*; *ġazāl* 'gazelle', Ar. *ġazāl*; *ġäläṭ*, *ḵäläṭ* 'error', Ar. *ġalaṭ*; *mäġrib*, *mäḵrib* 'evening', Ar. *maġrib*; *ġärīb*, *ḵärīb* 'stranger', Ar. *ġarīb*; *gaz* 'battle', Ar. *ġazw*.

ARGOBBA: *qärš* 'thaler', Ar. *ġirš*.

GURAGE: *baliq* (Ed.Ms.S.W.) 'old man', Ar. *bāliġ*; *qärši* (E.Ms.S.W.Z.) 'silver', Ar. *ġirš*; *gaz* (Gur.) 'battle', if it is to be compared with Ar. *ġazw* 'military expedition'.

In the few native roots with *ġ*, Arabic *ġ* corresponds to Geez, Tigre, Tigrinya ', South Ethiopic *zero*. Example: G. *'arbä* 'the sun sets' (Ar. *ġaraba*), Tna. *'aräbä*, Amh. *arräbä*.

In the Cushitic languages, the sound *ġ* of the Arabic loanwords is rendered by *q* and *g*, in Somali also by *ġ*. Thus, Sidamo: *baliqi* (Qabenna) 'old', Ar. *bāliġ*. Somali: *qado* 'midday meal', Ar. *ġadā'*; *šuqul* (and *šuġul*) 'work', Ar. *šuġl*; *qarīb* and *ġarīb* 'stranger', Ar. *ġarīb*. Bedja: *mugrāf* 'cup', Ar. *miġrafa*; *ragīf* 'bread', Ar. *raġīf*.[35]

Classical Arabic *ġ* appears as *q* in the Beduin dialects of Algeria, in some other places as *g*. Needless to say, none of the Arabic loanwords in Ethiopic come from any of these regions.[36]

[33] For Arabic *ġ* represented as *g* in native Geez words, cp. G. *šägärä* 'run', Ar. *šaġara*; G. *gime* 'fog' if it is to be compared with Ar. *ġaym*, as suggested by Brockelmann, *Grundriss*, p. 123. The G. *gime* is most probably of Cushitic origin as suggested by E. Cerulli, *Studi etiopici. IV. La lingua caffina*, p. 444.

[34] Note that *ġ* exists in Harari only in loanwords.

[35] Bedja has no *q*.

[36] Kampfmeyer, *op. cit.*

LARYNGEALS

ARABIC ' = Geez ', occasionally *zero* finally; Tigre, Tigriṅya '; Amharic *zero*; Harari *zero* initially, ' medially and finally; Argobba *zero*, '; Gurage *zero*.

GEEZ: *'abäbi* 'flood', Ar. *'ubāb*; *'adim* 'red hide', Ar. *'adīm*; *wäba* 'malaria', Ar. *wabā'*; *baqela* 'bean', Ar. *bāqilā'*.

TIGRE: *'əbrät* 'needle', Ar. *'ibra*; *'awwäl* 'formerly', Ar. *'awwal*.

TIGRIṄYA: *'abriq* 'jar', Ar. *'ibrīq*; *'əwan* 'time', Ar. *'awān*.

AMHARIC: *ard* 'earth', Ar. *'arḍ*; *aṭlas* 'black silk', Ar. *'aṭlas*.

HARARI: *əzni* 'permission', Ar. *'iḏn*; *ummät* 'community', Ar. *'umma*; *qara'a* 'read', Ar. *qara'a*; *mäla'ikat* 'angel', Ar. *malā'ikat* 'angels'; *lu'lu'* 'pearl', Ar. *lu'lu'*.

ARGOBBA: *qarra* 'read', Ar. *qara'a*; *abšər* 'courage!', Ar. *'abšir* 'bring good news!'; *rabi 'awwal* 'name of month', Ar. *rabī' 'awwal*.

GURIAGE: *ayam* (S.Z.) 'day', Ar. *'ayyām* 'days'; *bur* (M.Go.Z.) 'pit, well', Ar. *bi'r*; *däwi* (S.W.) 'medicament', *däwa* (Z.), Ar. *dawā'*.[37]

ARABIC ' = Geez ', later ';[38] Tigre, Tigriṅya '; Amharic *zero*; Harari *zero* initially, ' medially; Argobba *zero* initially, ' medially and finally; Gurage *zero*, '.

GEEZ: *'aḷarəd* 'Mercury', Ar. *'aṭārid*; *'angug* 'sea lizard', Ar. *'unǧūǧ*; *'äräbon*, *'äräbon* 'deposit', Ar. *'arabūn*; *qʷənza'l*, *qʷänza'l* 'curled hair', Ar. *qunzu'a*.

TIGRE: *'addäla* 'distribute', Ar. *'addala*; *ǧəm'al* 'Friday', Ar. *ǧum'a*.

TIGRIṄYA: *'allämä* 'instruct', Ar. *'allama*; *lä'abä* 'be tired', Ar. *la'iba*.

AMHARIC: *aqəl* 'reason', Ar. *'aql*; *äräqi* 'liquor', Ar. *'araq*; *mädf* 'canon', Ar. *midfa'*; *mänfa* 'utility', Ar. *manfa'a*.

[37] For Semitic ' represented occasionally as ' in Ennemor, and Zway, see *JAOS* 71 (1951). 214.

[38] For the confusion of the laryngeals in later stages of Geez, see Ullendorff, *op. cit.* 35ff.

HARARI: *išqi* 'love', Ar. *'išq*; *aqli* 'plan', Ar. *'aql*; *ma'ruf* 'favor', Ar. *ma'rūf*; *ǧum'a* 'Friday', Ar. *ǧum'a*.

ARGOBBA: *afiya* 'health', Ar. *'āfiya*; *abdəqadər* 'Wednesday', Ar. *'abdu-l-qādir*, proper noun; *ǧum'a* 'Friday', Ar. *ǧum'a*; *zil qi'da* 'name of month, Ar. *ḏū-l-qa'da*; *du'a mäññä* 'pray', Ar. *da'wā* 'prayer'.

GURAGE: *ibad* (S.W.) 'prayer', Ar. *'ibāda*; *umr* (Ms.Go.) 'age', Ar. *'umr*; *aynät* (M.Ms.A.S.W.) 'kind', Ar. *'aynat*; *ǧimat* (C.E.S.W.) 'Friday', *ǧimät* (Z.), *ǧim'at* (En.Gt.Ed.), Ar. *ǧum'a*.

ARABIC *h* = Geez *h* (also written later *ḥ*, *ḫ*); Tigre, Tigrinya *h*; Amharic *h*, [*zero*]; Harari *ḥ*, /*h*/; Gurage *zero*, [*h*]; Argobba, no examples.

GEEZ: *həlal* 'new moon', Ar. *hilāl*; *zäwhər* 'gem', Ar. *ǧauhar*; *'əlḥažra* 'the Hidjra', Ar. *'al-hiǧra*; *dərhəm*, *dərḥəm* 'drachma', Ar. *dirham*.

TIGRE: *dähäb* 'gold', Ar. *ḏahab*; *hadma* 'destroy', Ar. *hadama*.

TIGRINYA: *had'e* 'be quiet', Ar. *hada'a*; *rahya* 'pool', Ar. *rahw*.

AMHARIC: *qahwa* 'coffee', Ar. *qahwa*; *həwa* 'air', Ar. *hawā'*; *drim* 'drachma', Ar. *dirham*.

HARARI: *ḥawā'* 'air', Ar. *hawā'*; *miḥra* 'affair', Ar. *mihra*; *aḫli* 'tribe, clan', Ar. *'ahl*; *h* only in *zāhid* 'monk', Ar. *zāhid*; *šāhid* 'witness', Ar. *šāhid*; *qahwa* 'coffee', Ar. *qahwa*.

GURAGE: *qawa* 'coffee', Ar. *qahwa*; *alla* (S.W.) 'God', Ar. *'allāh*; *ṫahlil* (S.) 'prayer of death commemoration', Ar. *ṫahlīl* 'profession of faith'.

ARABIC *ḥ* = Geez *ḥ*, later *h*; Tigre *ḥ*; Tigrinya *ḥ*, (*h*); Amharic *zero*, *h*; Harari *ḥ*; Argobba /*ḥ*/, [*h*]; Gurage *zero*, *h* (according to the dialects).

GEEZ: *ḥadid* 'iron', Ar. *ḥadīd*; *ḥarir* 'silk', Ar. *ḥarīr*; *hafir* 'protector', Ar. *ḥafīr*; *rähim* 'merciful', Ar. *raḥīm*.

TIGRE: *ḥassäna* 'do good', Ar. *hassana*; *saḥəl* 'shore', Ar. *sāḥil*.

TIGRINYA: *ḥalal* 'legitimate', Ar. *ḥalāl*; *ḥarir* 'silk', Ar. *ḥarīr*; *huqqʷa* 'snuff box', Ar. *ḥuqqa*; *bä'al habäba* 'sympathetic person', Ar. *ḥabba* 'love'.

AMHARIC: *araǧ*, *haraǧ* 'auction', Ar. *ḥarāǧ*; *abäsuda* 'kind of mint', Ar. *ḥabba sawdā'* 'black grain'; *kʷəl* 'antimony', Ar. *kuḥl*; *hakim* 'physician', Ar. *ḥakīm*; *isab*, *həsab* 'count', Ar. *ḥisāb*.

HARARI: *ḥaraka* 'movement', Ar. *ḥaraka*; *ḥisti* 'jealousy', Ar. *ḥisd*; *saḥib* 'friend', Ar. *ṣāḥib*.

ARGOBBA: *märḥaba ala* 'agree', Ar. *marḥaba* 'welcome!'; *muḥarram* 'name of month', Ar. *muḥarram*; *harb* 'spear, battle', Ar. *ḥarb*.[39]

GURAGE: *amd* (C.En.) 'thanks', Ar. *ḥamd*; *aq* (W.Go.) 'truth', *haq* (S.), Ar. *ḥaqq*; *akim* (Z.) 'physician', *hakim* (A.S.), Ar. *ḥakīm*; *haǧa* (W.Ms.) 'affair', Ar. *ḥāǧa*.

In summing up, the correspondence between the Ethiopic and the Arabic phonemes is as follows. The parentheses ( ) indicate a rare correspondence; the brackets [ ] indicate an only example; the symbols / / indicate that the Ethiopic sound occurs only in Arabic loanwords; the symbols | | indicate that the Ethiopic sound is an internal or dialectal variant.

University of California, Los Angeles.

[39] Note that non-geminated *ḥ* occurs in Argobba only in Arabic loanwords.

ARABIC	GEEZ	TIGRE	TIGRINYA	AMHARIC	HARARI	ARGOBBA	GURAGE
ʾ	ʾ (initial), zero (final)	ʾ |	ʾ	zero	zero (initial) ʾ (medial, final)	zero, [ʾ]	zero
ʿ	ʿ, late ʾ	ʿ	ʿ	zero	zero (initial) ʾ (medial, final)	zéro (initial, fina) ʾ (medial)	zéro, [ʾ]
b	*b*	*b*	*b*	*b*	*b*	*b*	*b*
d	*d*	*d*	*d*	*d*	*d*	*d*	*d*
ḏ	*d*, *z*	*d*, *z*	*d*, *z*	*z*, [*d*], [*t*]	*z*	*z*	[*d*], [*z*]
ḍ	*d*, [*ṣ*]	*d*, [*ṣ*]	*d*	*d*, *ṭ*, [*z*]	*d*	*d*	...
f	*f*	*f*	*f*	*f*	*f*	*f*	*f*
ǧ	*g*, *g*^{w}, (late *ǧ*, *ẓ*)	*ǧ*, *g*	*ǧ*, *g*, *g*^{w}	*ǧ*, *g*, *g*^{w}	*ǧ*, *(g)*	*ǧ*	*ǧ*, *g*, *g*^{w}, |*g*ʿ|
ġ	*q*	*q*	*q*, *q̄* (intervocalic)	*q*, *q*^{w}, [*g*]	*ḵ*, *q*, /*ġ*/, [*g*]	[*q*]	*q*, [*g*]
h	*h*, (late *ḥ*, *ḫ*)	*h*	*h*	*h*, [zero]	*ḥ*, /*h*/	...	zero, [*h*]
ḥ	*ḥ*, (late *h*)	*ḥ*	*ḥ*, *(h)*	zero, *h*	*ḥ*	/*ḥ*/, [*h*]	zero, *h*
ḫ	*k*, *k*^{w}, *ḫ*	*k*, *(ḫ)*	*k*, *k*^{w}, *ḵ* (intervocalic), *ḫ*	*k*, *k*^{w}, *(h)*	*k*, *ḵ*, *(ḫ)*	*k*, *ḵ*	*k*, *ḵ*, *(h)*. [zero]
k	*k*, *k*^{w}	*k*	*k*, *k*^{w}, *ḵ* (intervocalic)	*k*, *k*^{w}	*k*, *(ḵ)*	*k*	*k*, *k*^{w}
l	*l*	*l*	*l*	*l*	*l*	*l*	*l*, |*n*|, |*r*|
m	*m*	*m*	*m*	*m*	*m*	*m*	*m*
n	*n*	*n*	*n*	*n*	*n*	*n*	*n*, |*r*|
q	*q*, *q*^{w}	*q*, *(g)*	*q*, *q*^{w}, (*q̄* intervocalic)	*q*, *q*^{w}	*q*, [*ḵ*]	*q*	*q*, [ʾ]
r	*r*	*r*	*r*	*r*	*r*	*r*	*r*, |*n*|
s	*s*, *(š)*	*s*, *(š)*	*s*, *(š)*	*s*, [*š*]	*s*	*s*	*s*
ṣ	*ṣ*, *s*	*s*, *(š)*, [*ṣ*]	*s*, *š*, *(ṣ)*	*s*, [*š*]	*s*	*s*, *s*^{w}	*s*
š	*š*, *(s)*	*š*	*š*	*š*	*š*	*š*	*š*
ṯ	*l*	*ṯ*	*ṯ*	*l*	*ṯ*	*ṯ*	*ṯ*
ṱ	*l*, *s*	*l*, [*s*]	*s*, [*t*]	[*s*], [*t*]	*s*, *t*	[*s*]	[*t*]
ṭ	*ṭ*	*ṭ*	*ṭ*	*ṭ*	*ṭ*	*ṭ*	*ṭ*
ẓ	[*t*]	*d*, *z*	[*t*]	[*d*]	*z*	...	[*z*]
w	*w*	*w*	*w*	*w*	*w*	*w*	*w*
y	*y*	*y*	*y*	*y*	*y*	*y*	*y*
z	*z*	*z*	*z*	*z*	*z*	*z*	*z*

CLASSIFIERS IN VIETNAMESE

NGUYEN ĐINH HOA

0. In the Vietnamese language, anything counted is expressed by a phrase consisting of the numeral and the designation of the thing counted, the latter being preceded by a modifying word of fairly material content and expressing the unit of the object-idea.

This modifying word belongs to a subclass of the word-class of nouns. It has been given different names by earlier French and Vietnamese students of the language : "appellatif, nom générique, nom numérique" (Pétrus Trương Vĩnh Ký, *Grammaire de la Langue Annamite*, Saigon, Guilland and Martinon, 1883), "numéral" (Nicholas Trương Vĩnh Tống, *Grammaire de la Langue Annamite*, Saigon, Đức-Lưu-Phương, 1932), "substantif générique" (A. Chéon, *Cours de Langue Annamite*, 2nd ed., Hanoi, F.-H. Schneider, 1904), "générique, classificateur" (Bouchet, *Cours Élémentaire d'Annamite*, Hanoi-Haiphong, Imprimerie d'Extrême-Orient, 1908), "pronominal, numéral" (Edward C. Chodzko, *Introduction à l'Étude de la Langue Annamite*, Haiphong, 1932), "spécificatif" (P. Francastel, *Précis de Langue Vietnamienne*, Saigon, Imprimerie de l'Union, 1948), "classificateur" (Lê Văn Lý, *Le Parler Vietnamien*, Paris, Hương-Anh, 1948), "article générique" (R. Bulteau, *Cours d'Annamite*, 3rd ed., Paris, Larose, 1950), "loại-từ" /classifier/" (Trần Trọng Kim, Bùi kỷ and Phạm Duy Khiêm, *Việt-Nam Văn-Phạm*, 3rd ed., Saigon, Tân-Việt, 1950), "classifier" (M. B. Emeneau, *Studies in Vietnamese (Annamese) Grammar*, Berkeley, University of California, 1951). The name "classifier" or "classificateur" seems to be most adequate, since it shows that Vietnamese speech-feeling divides the material universe into a number of classes which receive distinction in the numerative word. This is an example of classification by "congruence".[1] And since

[1] Leonard Bloomfield, *An Introduction to the Study of Language*. (New York: Henry Holt, 1914) pp. 130-131.

Vietnamese nouns do not have number as part of their class meaning, but merely serve to identify the species, the class meaning of classifiers is really "one unit quantity, or number, of that denoted by the noun that it precedes."[2]

Emeneau's statement on Vietnamese classifiers[3] is about the most illuminating to date, but, like previous writers, he fails to mention stress patterns involved in constructions containing a classifier. This paper proposes to remedy this oversight and also to present sub-classes of non-classified nouns (in 3) as well as to identify various sub-members of the sub-heading "special classifiers (in 6.2).

1.0. There is a considerable number of nouns falling under the sub-heading of classifiers in Vietnamese. The choice of these class-denoting words, very important syntactically, depends upon the nature of the thing counted, although the numerals are the same for all classes of objects counted. The word order for this kind of numeration is the following: numeral plus classifier plus classified noun. (The syllables are run together, and there is no pause or hesitation. Heavy stress is indicated by /ˊ/, light stress by /ˇ/, whereas a syllable with normal stress does not have any mark preceding it.)

Examples:

/ba·ˇká·y'bá·t/ *ba cái bát* (three+non-living thing+eating bowl) 'three eating bowls'

/bónˇkɔn'mɛ̀w/ *bốn con mèo* (four+living thing+cat) 'four cats'

/ha·yˇtə̀'bá·w/ *hai tờ báo* (two+sheet+newspaper) 'two newspapers'

/mấyˇkʌy'le/ *mấy cây lê* (how many+plant+pear) 'how many pear trees?'

/mấyˇkwả·'cuáy/ or /và·yˇkwả·'cuáy/ *mấy/vài quả chuối* (a few+fruit+banana) 'some bananas'

/namˇkʌn'thịt/ *năm cân thịt* (five+pound+meat) 'five pounds of meat'

These constructions suggest such English phrases as "five head of cattle," or "a bevy of larks (or quail)," "a clowder of cats," "a covey of partridges," "a drove of oxen," "a flock of sheep,"

[2] Murray B. Emeneau, *op. cit.*, p. 93.

[3] *Ibid.*, pp. 93-113.

"a herd of cattle," "a kendle of kittens," "a shoal (or school) of fish," etc.

When number is emphasized, the first syllable of the phrase takes heavy stress: /ba·ˇká·y'bá·t/ 'three (*not* four) eating bowls,' /'bónˇkɔn'mɛ̀w/ 'four (*not* two) cats.'

In literary constructions where inversion is necessary, the word order may become *noun plus numeral plus classifier*, as in these coupled lines of poetry:

/lɔmkhɔm zɯ́áynúy|tiáwvà·ycú/
/lá·kdá·k bensʌwŋ|cɛ̀mɤ́yñà·/
lom khom dưới núi, tiều vài chú
lác đác bên sông, chợ mấy nhà

'At the foot of the mountain a few bent-down woodsmen,
By the riverside some scattered market stalls.'

The orders $NU+CL+N$ and $N+NU+CL$ (NU, number, CL, classifier, N, noun) are the only ones possible: /và·ycútiàw/ *vài chú tiều* or /tiʌwvà·ycú/ *tiều vài chú* 'a few woodsmen.' The impossibility of $*NU+N+CL$, $*N+CL+NU$, $*CL+N+NU$ or $*CL+NU+N$ points out the inseparability of the pair *number* plus *classifier*, recently duly highlighted by Smalley in a course[4] published in tentative form after the submission of this paper. Smalley uses (pp. 122-125) the term "countable word", whose ingenious definition reads (123) "any word which may directly follow một (or any number from 1-9) in an isolated phrase consisting of no more than the two words in question."

The order $N+NU+CL$ is regularly found in Thai, Lao and Burmese, although in the latter case, if the numeral is above ten and a multiple of ten, the arrangement becomes

$N+ʔaCL+NU$

lú ʔayauʔ hnashé	'twenty persons' (yauʔ 'human being')
myîn ʔakáun Oộunzé	'thirty horses' (káun 'animal')

1.1. There are two general classifiers that are found frequently: *cái* and *con*, used before nouns denoting inanimate things and

[4] Smalley, William A. and Nguyên Văn Vận, *Tiếng Việt Cho Các Giáo-Sĩ, Vietnamese for Missionaries. A course in the spoken and written language of Central Vietnam.* Trial Edition. Dalat, Vietnam: Imprimerie Evangélique, 1954. This work has been reviewed by me in *Word*, 12.1 (1956). 167-170.

nouns denoting non-human animate things, respectively. The use of these general classifiers, however, is subject to several exceptions. There are, for instance, some nouns denoting small living things which are classified with *cái*: *kiến* 'ant', *ong* 'bee,' etc. Conversely, *con* can also be used both for nouns denoting inanimate things, e.g. *dao* 'knife', *đê* 'dike,' *đường* 'road,' *thuyền* 'boat', etc., and for nouns denoting human beings, as in *con bạc* 'gambler,' *con người* 'human being,' etc.

1.2. Besides *con* and *cái*, there exist some two hundred special classifiers which students of the language just have to memorize in connection with different nouns. It is true that a certain amount of freedom is allowed in the use of *cái*—the latter indeed encroaches upon the special classifiers without making much difference. 'A receipt' may be *một lờ biên-lai* (*lờ* being the classifier for nouns denoting sheets of paper, metal, etc.) or just *một cái biên-lai*. Another example is *một thanh kiếm* (thanh 'thin sheet of material'), or *một lưỡi kiếm* (lưỡi 'blade'), or merely *một cái kiếm* for 'a sabre'. Some native speakers even prefer to use *cái* in all instances. This freedom, however, has certain limits, and each subclass of classified nouns requires its own classifier or classifiers, as we shall see in 6.2.

Following are only some examples to show how the meaning changes with the classifier:

đào 'peach': *cây đào* 'peach tree', *cành đào* 'branch of peach tree', *hoa đào* 'peach blossom', *quả đào* 'peach (the fruit)'.

chè 'tea': *cây chè* 'tea shrub', *lá chè* 'tea leaf', *nụ chè* 'tea bud', *búp chè* 'unfolded tea leaf', *hạt chè* 'tea bud', *nõn chè* 'unfolded tea leaf', *nương chè* 'tea grove', *bã chè* 'tea dregs', *ấm chè* 'teapot', *bình chè* 'teapot', *nước chè* 'tea (the beverage)'.

báo 'newspaper': *số báo* 'issue, number, copy of newspaper', *trang báo* 'page in newspaper', *mặt báo* 'page in newspaper', *cột báo* 'column in newspaper', *bài báo* 'newspaper article'.

It is clear then that a good dictionary of the Vietnamese language should give all instances of usage for each classifier, and also note each time the classifier to be used with each classified noun.

1.3. When the material thing designated is not discrete, the classifier indicates *quantity* rather than *number*. Examples.

hai cân thịt 'two pounds of meat,' *ba bát cơm* 'three bowl(ful)s of rice,' *năm thước vải* 'five meters of cloth,' etc.

Cân, *bát* and *thước* here denote units of measurement or 'contents', and *nọt* instruments of measurement or 'containers,' in which case they would require general classifier *cái*: *một cái cân* 'a balance, a pair of scales,' *hai cái bát* 'two eating bowls,' *một cái thước* 'a yardstick, meter stick, or ruler.'

Other examples are:

Names of containers or measures[5] (classified with *cái*)		*Classifiers showing contents*
ấm	'kettle, teapot, coffeepot'	'kettleful, potful' e.g., *một ấm chè* 'a potful of tea.'
bình	'teapot, coffeepot'	'potful' e.g. *một bình càphê* 'a potful of coffee.'
bị	'woven-reed bag with handle'	'bagful,' e.g. *bốn bị thóc* 'four bagfuls of paddy.'
bì	'sackcloth bag'	'bagful,' e.g. *ba bì muối* 'three bagfuls of salt.'
thìa	'spoon'	'spoonful,' e.g. *hai thìa đường* 'two spoonfuls of sugar.'

1.4. Some ordinary nouns are used as classifiers on a temporary basis. For example, in một cái bàn 'a table,' bàn is a noun classified with the general "numerary adjunct" or "auxiliary noun." But it can in turn be used as a temporary classifier: *một bàn tiệc* 'a tableful of food (in a banquet).'

Similarly, we have:

một cái nhà	'a house'
một nhà rác	'a houseful of refuse or wastepaper.'

1.5. *Cái* and some other classifiers denoting specific actions occur as objects of functive verbs:[6] the construction verb plus numeral plus classifier specifies the number of times the action is performed.[7] Examples:

[5] For the use of the word «measure» to designate classifiers in Chinese for instance, see M. Gardner Tewksbury, *Speak Chinese* (New Haven: Yale University Press, 1948) p. 17, John De Francis, *Beginning Chinese* (New Haven: Yale University Press, 1946) pp. 31-32.

[6] A functive verb predicates action or event whereas a stative verb describes a quality or a condition. A number of auxiliary verbs such as *rất* 'very' occur only before a stative verb, never before a functive verb.

[7] This type of classifier which indicates single action is sometimes called «semelfactive classifier.» Used after a verb of striking, hitting, kicking and the like, it means 'once,' 'twice,' etc. depending on the numeral. When the term specifies the weapon with which the blow is delivered, the usual translation is 'a blow with.'

ngủ một giấc/cái	'to take a nap'
tắm một cái	'to take a bath'
tát hai cái	'to give two slaps'
vụt một trận	'to give a whipping'

2.0. Besides the substantive phrase mentioned in 1.0. and consisting of a numeral, a classifier and a classified noun, there are other constructions where a classifier enters.

2.1. The classifier and the classified noun may be followed by an indicator, interrogative or demonstrative. Examples:

con mèo nào?	'which cat?'
con mèo này	'this cat'
con mèo ấy	'that cat'
con mèo kia	'that cat over there, the other cat'

2.2. In less precise speech the use of the classifier is optional. People say:

nhà này or *cái nhà này*	'this house'
làng ấy or *cái làng ấy*	'that village'

2.3. Attributes showing order consist of word *thứ* 'kind, order, rank' plus a cardinal number[8] or *mấy* (or *bao nhiêu*) 'how much/how many.' When the classified noun is followed by such an attribute which refers to a specific unit, the classifier may be omitted too:

(*cái*) *nhà thứ mấy*?	'which house? (which in order?)'
(*cái*) *nhà thứ hai*	'the second house'

2.4. The phrase may finally contain an attribute. The order then is classifier plus noun plus attribute plus indicator (if any):

con mèo đen	'the black cat'
con mèo đen ấy	'that black cat'

2.5. The phrase may contain both a numeral at the beginning and an indicator at the end, when the species identified by the noun is fully numerated, that is, both counted and pointed at. Examples:

[8] The following substitutions are made within the first decade (1-10): *nhất* (<Chinese *yi*) for một 'one,' *tư* (Sino-Vietnamese *tử* <Chinese *sì*) for *bốn* 'four,' and optionally *nhì* (Sino-Vietnamese *nhị* <Chinese *èr*) for *hai* 'two.'

hai con mèo nào? 'which two cats?'
hai con mèo này 'these two cats'
hai con mèo ấy 'those two cats'
hai con mèo kia 'those two cats over there, the other two cats'

2.6. The classifier may also be used as substitute when the antecedent has already been sufficiently identified by the previous context:

Tôi có ba con mèo, hai con trắng, một con đen
'I have three cats, two white and one black.,
Tôi có hai quyển sách, một quyển mỏng, một quyển dầy
'I have two books, one thin and one thick'
Ông muốn mua con nào?—con bé
'Which (animal) do you want to buy?—'the small one'
Bà muốn mua mấy quả?—ba quả
'How many (fruits) do you want to buy?'—'Three'
Con ấy là con gì?—Con ấy là con ba ba
'What's that?'—'That's a turtle.'

2.7. When the noun is identified by an attribute consisting of a full predication introduced optionally by *mà* (which as a coordinating conjunction means 'but strangely', and as a final particle emphasizes the content of a predication), the appropriate classifier is used before the noun. *Mà* then can be translated by the relative pronoun 'which' or 'whom'. Examples:

cái bàn (mà) tôi mua tháng trước
'the table (which) I bought last month'
con voi (mà) anh thấy lúc nãy
'the elephant (which) you saw a while ago'

Some people use *cái* before the appropriate classifier (but not before *cái* itself) or a nonclassified noun (see 3.):

cái con voi (mà) anh thấy lúc nãy
'the elephant (which) you saw a while ago'
cái người (mà) tôi vừa chào
'the person (whom) I just said "hello" to'
cái độc-lập (mà) chúng tôi muốn
'the (kind of) independence (that) we want'

However, this use of the "extra" *cái*[9] is considered to be old-fashioned. Indeed, a speech which is interspersed with classifier *cái* preceding such abstract terms as *độc-lập* 'independence', *tự-do* 'freedom', *bình-đẳng* 'equality' (see 3.7.) is likely to strike young people as being "incorrect".

Another construction may occur where the attribute predication, introduced by *mà*, has the pronoun *nó'he*, 'she (of children, adults of low status), it' as "extra" subject besides the noun in front.

con voi (*mà*) *nó dậm chết thằng hề.*
'the elephant (which) trampled the clown to death'
thằng bé (*mà*) *nó vừa chào tôi*
'the boy who just greeted me'

This construction can also be found frequently in a complete sentence:

cái thằng chồng em nó chẳng ra gì
'that husband of mine, he isn't worth anything'

It can then be best analyzed as consisting of a normal predication *nó chẳng ra gì*, preceded by another one, *cái thằng chồng em*, a substantive phrase "announcing" the second predication and followed by a comma pause.

3. Beside nouns which require a classifier, there are several nouns which do not. Nonclassified nouns include a number of sub-classes. Following is a tentative list of some of them:

3.1. Nouns denoting substance matter (*chất*), color (*mầu*, *sắc*), smell (*mùi*), taste (*vị*), noise (*tiếng*), etc.

3.2. Nouns denoting time units:[10] *buổi* 'half a day, session', *bữa*, 'day, meal', *canh* 'nightwatch', *chốc* 'short moment', *dạo* 'period, time', *đời* 'lifetime, generation, period', *đêm* 'night', *độ* 'period, time', *giây* 'second', *giờ* 'hour', *hôm* 'day', *khắc* 'quarter-hour', *khi* 'time—when something happens', *kỳ* 'session, time', *kỷ-nguyên* 'era', *lần* 'time, round', *lát* 'short moment', *lúc* 'moment', *mùa* 'season', *năm* 'year', *ngày* 'day', *phút* 'minute', *tháng* 'month',

[9] Maurice Grammont and Lê Quang Trinh, « Études sur la Langue Annamite,» *Mémoires de la Société Linguistique de Paris*, 17. (1911-1912). 204. Trương Vĩnh Tông, *op. cit.* pp. 23ff, called this "extra" *cái* a "definite article", and Bulteau, *op. cit.*, p. 21, called it a "sort of super-article".

[10] These are often used as classifiers, and have been called "quasi-classifiers."

thế-kỷ 'century', *thời-đại* 'period', *thời-kỳ* 'period', *tuần* (*lễ*) 'week', *tuổi* 'year of age', etc.

3.3. Nouns denoting geographical areas, regions, or administrative or social units:[11]

bộ 'one of the three sections of Vietnam', *chỗ* 'place', *chôn* 'place', *huyện* 'district', *khu* 'area', *kỳ* 'one of the three sections of Vietnam', *làng* 'village', *miền* 'area, region', *nơi* 'place, location', *phần* 'part, one of the three sections of Vietnam', *phủ* 'prefecture', *thôn* 'hamlet', *tỉnh* 'province', *tổng* 'canton', *vùng* 'area', *xã* 'commune', *xóm* 'hamlet', etc.

3.4. Nouns denoting kinds, grades, classes of things, groups of people, including army units:

giống 'species, breed, race', *loài* 'species', *loại* 'kind', *thứ* 'kind', *bè* 'faction', *bọn* 'group, gang', *đám* 'crowd', *đảng* 'party', *lũ* 'band', *nhóm* 'group', *phái* 'faction, party', *tụi* 'gang', *tốp* 'small group, squad', etc.

3.5. Several additive constructions using two nouns denoting synonymous or related items :

nhà cửa 'dwelling, house'	(lit.: house and door)
giầy dép 'footwear'	(lit.: shoes and sandals)
chăn chiếu 'bed clothing'	(lit.: blanket and mat)
sách vở 'books and supplies'	(lit.: books and notebooks)

3.6. Compound nouns—most of which are Chinese loans—such as:

độc-lập 'independence', *ảnh-hưởng* 'influence', *hạnh-phúc* 'happiness', *kết-quả* 'result', *tự-do* 'liberty', etc. However, the nouns denoting concepts or institutions, i.e. *dân-chủ* 'democracy', *tự-trị* 'self-government', *thông-nhất* 'unity', *phúc-lợi* 'welfare', etc., may take the classifier *nền* 'foundation, basis'.

4.0. *Classifiers for Nouns Denoting Human Beings.*

These classifiers belong to two categories: they may be general or special. The former do not indicate any status, and the latter do, that is to say, their use is governed by the "status factor",

[11] These nouns are sometimes called "autonomous classifiers" since they carry a full meaning when not followed by a head noun.

that is the social distance between the speaker and the person referred to.

4.1. *General Classifiers for Nouns Denoting Human Beings.*

4.1.1. *người* is used before kinship terms and other nouns giving an idea of the occupation of the person. Here are some examples :

một người cha	'a father'
một người vợ	'a wife'
một người bạn	'a friend'
một người khách	'a guest, visitor'
một người phu xe	'a rickshaw man'
một người thợ	'a worker'
một người bếp	'a cook'

người can also be followed by a stative verb or a functive verb used as attribute to show age, status, physical features, etc. Examples :

một người già	'an old person'
một người câm	'a mute person'
một người điếc	'a deaf man'
một người lạ	'a stranger'
một người ở	'a servant'
một người nghèo	'a poor man'

If employed without the numerator, the above phrases take a collective denotation, and mean respectively 'the old (people)', 'the mute (people)', 'the deaf (people)', 'the strangers', 'the servants', 'the poor (people)'.

4.1.2. *Kẻ*. *Người* can be replaced by *kẻ*, but not always. In fact the meaning of *kẻ* is a little vaguer than that of *người*. Besides, when the word which it precedes is not a verb, *kẻ* rarely substitutes for *người*. People do not say, for instance, **kẻ bếp*, **kẻ thợ mộc*, **kẻ phu xe*, etc., but *người bếp* 'the cook', *người thợ mộc* 'the carpenter', *người phu xe* 'the rickshaw man', etc., *bếp*, *thợ and phu* being nouns denoting respectively 'kitchen', 'artisan, worker', 'coolie'.

When the attribute following the classifier is a verb *kẻ* is used interchangeably with *người*. Examples:

kẻ/người ăn mày	'the beggar or beggars'
kẻ/người nghèo	'the poor people'

However, in certain expressions such as the following, *kẻ* cannot be replaced by *người*:

kẻ cắp	'the thief(s)'
kẻ trộm	'the robber(s)'
kẻ cướp	'the brigand(s), pirate(s), bandit(s)'
kẻ chợ	'city dweller'
kẻ sĩ	'scholar'

For literary effects, especially to avoid repetition or to follow the rules of parallelism,[12] the two are used in correlation:

kẻ giàu người nghèo	'some (are) wealthy and others (are) poor; the rich and the poor'
kẻ ở người đi	'some stayed and others left'
kẻ gậy người dao	'some (are) armed with sticks and others with knives'

4.1. *Chàng.* This classifier precedes some nouns denoting 'young men (of rather reputable character)'.

chàng thiếu-niên	'the young man'
chàng thanh-niên	'the young man'
chàng quân-tử	'the honest young man'
chàng rể	'the son-in-law'

Chàng can also take classifier *anh* or *chị* (See 4.2.1.3.):

anh chàng	'the fellow, the lad'
chị chàng	'the girl, the lass'

4.1.4. *Nàng.* This classifier is the female counterpart of *chàng.* Both of them appear to be used in speech nowadays less often than formerly.

nàng tiên	'immortal goddess, fairy'
nàng dâu	'daughter-in-law'
nàng hầu	'concubine'

[12] The tone of *người* belongs to the «*bằng*» (or even) class, comprising the low falling and high level tones, whereas that of *kẻ* belongs to the «*trắc*» (or uneven) class which includes the four other tones, high rising, low level, creaky level, creaky falling. These tones, so named by Emeneau, *op. cit.* p. 8, have the Vietnamese names *huyền, bằng, sắc, nặng, ngã* and *hỏi*, respectively. For an acoustic treatment of tones, see also Lê Văn Ly's *Le Parler Vietnamien*, cited in O. pp. 106-117. On parallelism or the coupling of lines—a *bằng* tone opposing a *trắc* tone—see Nguyen Đình Hòa, «Double Puns in Vietnamese—A case of 'Linguistic Play'», WORD 11 (August 1955), 237-244.

Both *nàng dâu* and *nàng hầu* may take the classifier *người* (4.1.1.) or *cô* (4.2.13.).

When *nàng* is preceded by classifier *cô* or *bà*, the expression takes a slightly different meaning:

cô nàng	'the damsel',
bà nàng	'the Goddess'.

4.2. Special Classifiers for Nouns Denoting Human Beings. These indicate the status of, and the degree of familiarity and intimacy toward, the person referred to. They include two subclasses, one of which is made up of kinship terms.

4.2.1. *Classifiers which are homonymous with kinship terms.*

4.2.1.0. Most terms of relationship are used as classifiers for nouns denoting people who are not consanguineal relatives of the speaker, except that terms for 'uncles, brothers' and 'aunts, sisters'—ordinarily classified with *người* (4.1.2.)—can also be classified with *ông* 'father of EGO's father or mother' and *bà* 'mother of either parent', respectively:

một ông bác	'an uncle—paternal, older than one's father'
một bà dì	'an aunt—maternal, younger than one's mother'

Used before a personal name the word *ông* is translated 'Mr'.

4.2.1.1.a. *Ông* is used as classifier for nouns denoting "men of about thirty years of age, maybe poor, but respected, because of social position". Examples:

ông chủ nhà	'landlord'
ông lão	'old man'
ông lão ăn mày	'old beggar'
ông già	'old man'
ông lý-trưởng	'village mayor'
ông đốc	'doctor' (*đốc* < French docteur)
ông cô đạo	'priest'
ông cha[13]	'(Catholic) father'
ông giáo	'teacher'
ông đồ	'scholar, teacher'
ông vua	'king'

[13] Not to be confused with *người cha* 'the father' where *cha* takes general classifier *người* (4.1.2.).

ông cử	'M.A.'
ông tú	'B.A.'
ông nghè	'Ph.D.'

4.2.1.1.b. In proverbs, sayings, folk literature, and other popular utterances, the classifier *ông* is used instead of *con* before nouns denoting personified or deified animals or elements of Nature. Examples:

ông ba mươi	'the Lord Tiger'
ông bình vôi	'the Lime Pot Spirit'
ông cá	'the whale'
ông công	'the Kitchen God'
ông cọp	'the Lord Tiger'
ông giăng	'the Moon God'
ông giời	'Heaven'
ông hổ	'the Lord Tiger'
ông hùm	'the Lord Tiger'
ông mãnh	'the Evil Spirit (especially of somebody who had died young and without descendants)'
ông táo	'the Kitchen God'
ông tiên	'the immortal man'
ông tạo	'the Creator'
ông tơ	'the Matchmaker'
ông xanh	'Heaven'
ông voi	'the elephant'

4.2.1.2. *Bà* is the kinship term designating 'the mother of EGO's father or mother'. It is used as classifier before nouns denoting women of a certain age, deities and goddesses.

(a) *Bà* is the female counterpart of *ông*. It occurs in such phrases as these:

bà chủ nhà	'landlady'
bà lão	'old lady'
bà lão ăn mày	'old woman beggar'
bà đỡ	'midwife'
bà cốt	'sorceress'
bà già	'old lady'
bà phước	'Catholic sister'
bà đồng	'medium'
bà mụ	'midwife'
bà lắm	'midwife'
bà hoàng-hậu	'queen'

(b) A *bà giáo* is not necessarily a school mistress herself, but simply the wife of a school teacher. The ambiguity does not exist since the academic title, or the public office—and more recently the occupation—which it opens to is clearly that of a man:

bà tú	'the wife of a B.A.'
bà cử	'the wife of a M.A.'
bà đốc	'the doctor's wife'
bà tham	'the clerk's wife'
bà phán	'the interpreter's wife'
bà phủ	'the wife of a district chief'
bà huyện	'the wife of a prefecture chief'

4.2.1.3. The terms *bác*, *chú*, *cô*, *anh*, and *chị*, meaning respectively 'father's elder brother,' 'father's younger brother,' 'father's sister,' 'elder brother,' 'elder sister,' are used for people likely of rural origin, surely of lowly trade—and therefore poor—but always honest. Examples:

bác thợ cạo	'barber'
chú tiểu	'Buddhist novice'
chú lính	'army private'
cô lái đò	'ferryboat girl'
anh thợ rèn	'blacksmith'
chị hàng rau	'vegetable girl'

(a) *Chú* is also used to refer to certain alien residents, especially of Chinese or Indian nationality:

chú khách	'Chinaman'
chú tàu	id.
chú chiệc	id.
chú tây đen	'Indian'

(b) Personified animals in, say, tales for children, may take classifier *chú* or *bác*, *chú mèo mà chèo cây cau, hỏi thăm chú chuột đi đâu vắng nhà* 'Uncle Cat climbed up the areca palm tree to inquire where Uncle Mouse had gone away to.'

4.2.1.4. *Cậu* and *cô* are used for young people (students, usually young and unmarried, or members of some liberal professions recently opened to both sexes) considered as belonging to families of mandarins or other officials. Examples:

cậu ấm	'mandarin's son'
cậu bé	'little boy'

cậu học trò	'young student'
cô chiêu	'mandarin's daughter'
cô giáo	'young lady teacher, or teacher's wife'
cô đỡ	'young midwife'
cô hồn	'medium'

4.2.1.5. The term *thím* 'wife of the uncle who is one's father's younger brother' is reserved for Chinese women:

một thím khách	'a Chinese woman'
một thím sẩm	id.

4.2.1.6. *Thày* is originally a noun meaning 'master, teacher,' but borrowed into the kinship terminology, it substitutes for *cha* 'father.' As a classifier *thày* precedes nouns denoting young but educated men having a rather highly regarded social position such as that of white-collar workers:

thày đồ	'scholar (versed in Sino-Vietnamese letters)'
thày thông	'interpreter'
thày ký	'clerk'
thày lý	'village mayor'
thày bói	'blind fortune-teller'
thày tu or *thày dòng*	'religious man'
thày kiện	'lawyer'
thày thuốc	'medicine man'
thày lang	'healer'
thày chùa	'Buddhist monk'
thày giáo	'teacher'
thày địa-lý or *thày phong-thủy*	'geomancian'
thày phù-thủy	'sorcerer'
thày lướng	'fortune teller'

The word *thày* 'master' is also used as an honorific prefix added to names of saints, philosophers, sages, e.g. *Thày Mạnh* 'Mencius' *thày Tăng-tử* 'Tseng Tze'.

4.2.1.7. *Cụ* is the kinship term applied to 'the father or mother of one's grandparent.' It is used as classifier for nouns denoting very old people; *ông* or *bà* is the extra classifier used here to add the sex distinction: *một cụ già* may be either an old man or an old

woman, but there is no ambiguity in such phrases as *một ông cụ già* or *một bà cụ già*. A few other examples:

cụ cô	'very old man or woman'
cụ lớn	'old mandarin'
cụ đạo	'old Catholic priest'

4.2.1.8. *Em* 'younger sibling' is used to classify 'young children of school age' in a very affectionate sense:

một em nhỏ 'a little child'

4.2.2. *Classifiers Which Are Not Kinship Terms.*

4.2.2.0. Some of these classifiers are nouns having a derogatory meaning. Others refer to officials, or eminent characters.

4.2.2.1. *Thằng* and *con* are used before nouns denoting very young boys and girls, respectively, or boys and girls, respectively, of low social status, domestic servants for instance:

thằng bé	'the little boy'
con bé	'the little girl'
thằng chăn trâu	'the buffalo boy'
thằng mõ	'the village crier'
thằng xe	'the rickshaw boy (in private homes)'
con sen	'maid'
con ở	'maid servant'

Most of the time, however, *thằng*[14] and *con*[14] carry a strong derogatory sense as in the following examples:

thằng hề	'the clown, the farceur'
thằng lùn	'the dwarf'
thằng bần	'poverty'
thằng móc túi	'the pickpocket'
con đĩ	'the harlot'
con điên	'the mad wench'

[14] Trần Trọng Kim pointed out the "age factor" in the status system, when he said *(op. cit.* p. 50) that even mendicants are respectfully referred to by expressions classified with *ông* and *bà*:

ông lão ăn mày	'the old beggar'
bà lão ăn mày	'the old woman beggar'

Thằng cha and *con mẹ* are very strong expressions preceding the personal names of people toward whom contempt is expressed. They both can be preceded by "extra" classifier *cái*.

4.2.2.2. *Ả* is the classifier for nouns denoting beautiful female creatures associated with sex and opium.

ả giang-hồ	'prostitute'
ả đào/đầu	'songstress'
ả phù dung	'patron saint of opium smokers'

4.2.2.3. *Bọn* 'group of people' Examples:

bọn cướp	'the brigands'
bọn bảo-hoàng	'the monarchists'
bọn cộng-sản	'the communists'

4.2.2.4. *Bực/bậc* 'hero, great personage.' (The noun *bậc* 'rung (of ladder), step, rank'). Examples:

bậc anh-hùng	'hero'
bậc vĩ-nhân	'great man'
bậc anh-thư	'heroine'

4.2.2.5. *Đồ* 'individual, "espèce de."' Examples:

đồ bất-lương	'crook'
đồ vũ-phu	'brute'
đồ du-côn	'thug'

4.2.2.6. *Đứa* is the regular classifier for nouns denoting 'young persons of unspecified sex.' This explains their use before nouns denoting blood relatives younger than EGO—those nouns carrying no sex distinctions—namely *em* 'younger sibling,' *con* 'child,' *cháu* 'grandchild, nephew, niece.' Other examples:

đứa bé	'child'
đứa trẻ	'child'
đứa ở	'house servant'

4.2.2.7. *Đức* 'king, deity (of Buddhism and Confucianism).' Examples:

đức vua	'the king'
đức Phật	'Buddha'
đức Khổng-Tử	'Confucius'

4.2.2.8. *Đứng/đấng* 'hero, monarch, or being with supernatural power.' Examples:

đấng anh-hùng	'hero'
đấng anh-quân	'great King'
đấng Tạo-Hóa	'Creator'

4.2.2.9. *Gã* 'individual' Examples:

gã liều-phu	'woodcutter'
gã Sở-Khanh	'Don Juan'

4.2.2.10. *Mụ* 'contemptible woman, not necessarily old.' Examples:

mụ già	'old woman'
mụ dâu	'madam'
mụ Tú-bà	'madam'

4.2.2.11. *Nhà* 'expert.' Examples:

nhà chính-trị	'politician, statesman'
nhà báo	'newspaperman'
nhà du-lịch	'tourist'
nhà văn	'writer'
nhà buôn	'merchant'

4.2.2.12. *Quân* 'thief, robber, and the like.' Examples:

quân trộm cắp	'robber'
quân giặc	'bandit, rebel, invader'
quân du-côn	'ruffian'

4.2.2.13. *Tay* 'expert.' Examples:

tay mỹ-thuật	'artist'
tay lão-luyện	'expert'
tay anh-tài	'talented person'

4.2.2.14. *Tên* 'individual (soldiers, criminals, etc.).' Examples:

tên gian-phi	'crook, robber'
tên sát-nhân	'assassin'
tên tù	'convict'
tên quá-khích	'extremist'
tên khủng-bố	'terrorist'

4.2.2.15. *Trang* 'hero, outstanding character.'

trang hào-kiệt	'gentleman'
trang anh-hùng	'hero'
trang hảo-hán	'gentleman, hero'

4.2.2.16. *Vị* 'officeholder, outstanding character, awesome being (god, angel).'

vị chủ-tịch	'chairman'
vị thủ-tướng	'Premier'
vị anh-hùng	'hero'
vị thần	'god'

4.2.2.17. *Viên* 'officeholder, agent.'

viên lý-trưởng	'village mayor'
viên bí-thư	'secretary'
viên đội	'sergeant'
viên thủ-quĩ	'treasurer'

5. *Classifiers for Nouns Denoting Non-human Living Things.*

5.0. The classifier *con* is used before nearly all numerated nouns denoting "species of living creatures which are not human." Examples:

một con bò	'a (or, one) cow'
hai con mèo	'two cats'

5.1. It has been said, in 1. 1, that at times, *con* serves as classifier for nouns denoting inanimate things, too. The list of examples is not very long:

con bài	'card (in games)'
con chèo	'oar'
con cò	'postage stamp'
con cờ	'chessman'
con dao	'knife'
con dấu	'seal, stamp'
con đê	'dike'
con đường	'road, street'
con mắt	'eye'
con ngươi	'pupil (of eye)'
con nước	'tide'
con mực	'stick of Chinese ink'

con quay	'(spinning) top'
con sào	'pole (for punting)'
con số	'digit, figure, number'
con sông	'river'
con súc sắc	'dice'
con tầu	'boat, ship'
con tem	'postage stamp' (French timbre-poste)
con thò lò	'teetotum'
con thuyền	'boat (rowboat or sailboat)'
con thoi	'shuttle'
con tơ	'hank or skein of silk'

5.2. Classifier *con* also occurs in a few phrases denoting human beings of unspecified sex (1.1). For instance:

con bạc	'gambler'
con buôn	'merchant'
con hát	'actor, actress'
con tin	'hostage'

The explanation which suggests that « there is probably some derogatory connotation in the occurrence of *con*"[15] in such instances is not satisfactory in all cases. Nor is it applicable to this example: *con người* 'human being'.

5.3. Whereas ghosts *(ma)*, monsters *(quái)*, devils *(quỷ)* and the like which are believed to haunt every corner of the land of Vietnam are considered as "non-human living things" and classified accordingly with *con*, names of wild animals such as elephants, tigers, etc. take the classifier *ông*. (See the examples given in 4.2.1.1.b). The tiger, especially, is the most venerated or sacred animal and is the object of many superstitions and traditions. "He is never mentioned without his title (my lord Tiger) and, within the forest, the coolie will no longer mention him at all."[16]

6. *Classifiers for Nouns Denoting Inanimate Things.*

Besides the general classifier *cái*, there are several specific ones for nouns denoting nonliving things. We shall discuss *cái* briefly

[15] Emeneau, *op. cit.* p. 102.

[16] Gabrielle Vassal, *On and Off Duty in Annam*, (London: William Heinemann, 1910) p. 122. In this connection, M. B. Emeneau, in "Taboos on Animal Names," *Language* 24.1. 56-63, suggested that the hypothesis of mere avoidance by hunters is not satisfactory, and that a religious taboo is probably involved.

in 6.1, then list in alphabetical order, the most frequently encountered special classifiers together with typical examples where they occur.

6.1. The general classifier which comes before nouns denoting inanimate things is *cái*, e.g. *một cái bàn* 'a table,' *hai cái bút* 'two pens,' etc. However, as we have pointed out in 1.1, nouns denoting small insects can be classified with *cái*, as well as with *con*. Examples:

cái kiến	'ant'
cái ong	'bee'
cái chấy	'(head) louse'
cái ghẻ	'acarid'
cái rận	'(body) louse'

6.2. Following is a list, far from being complete, of the specific classifiers which are most frequently encountered. The symbols P, M, T, V, and Q stand respectively for "classifier proper," "measure word," "temporary measure word," "classifier for verbs," and "quasi-classifier." A "temporary measure word" is a noun which ordinarily has its own classifier (1.3). "Classifiers for verbs" are discussed briefly in 1.5, and "quasi-classifiers" consist of the small subclass of nonclassified nouns already discussed in 3.

ấm	T	a kettleful (of water: *ấm nước*), a potful (of tea: *ấm chè*, or coffee: *ấm cà phê*)
ang	T	a glass container (of goldfish: *ang cá vàng*)
áng	P	a cloud: *áng mây*
bài	P	a lesson (*bài học*), a song (*bài hát*), a speech (*bài diễn-văn*), etc.
bãi	P	a flat expanse, stretch (of grass: *bãi cỏ*; sand: *bãi cát*), beach (*bãi bể*); a spit (of betel saliva: *bãi quệt trầu*)
bàn	T	a tableful (of food: *bàn tiệc*); a hand (*bàn tay*), a foot (*bàn chân*) or a leaf-shaped object (iron for pressing clothes: *bàn là*; pedal: *bàn đạp*)
bản	P	a script (song: *bản hát/ca*; musical composition: *bản nhạc*; report: *bản báo-cáo*; resolution: *bản quyết-nghị*)
bao	T	a bagful, container, package (of rice: *bao gạo*; of tea: *bao chè*; of cigarettes: *bao thuốc lá*)
bát	T	an eating-bowlful (of noodles: *bát phở*)

bậc/bực[17]	P	a rank, category, step (of stairs), or rung (of ladder): *bậc thang*
bầu	T/P	a gourdful (atmosphere: *bầu không khí*; blood: *bầu máu nóng* or *nhiệt-huyết*)
bầy	P	a flock, herd (of hunting dogs: *bầy chó săn*)
bể	T	a tankful (of rain water: *bể nước mưa*)
bên	Q	a side, party (in a lawsuit) (plaintiff: *bên nguyên*; defendant: *bên bị*)
bìa	P	a square cake (of beancurd: *bìa dậu*)
bó	P	a bunch, bundle (of twigs: *bó củi*)
bộ	P	a set (suit of clothes: *bộ quần áo*; furniture set: *bộ bàn ghế*)
bụi	P	a bush, thicket (of bamboos: *bụi tre*; of trees: *bụi cây*, etc.)
bụng	T	a bellyful (of glutinous rice: *bụng cơm nếp*)
buồng	P	a bunch (of bananas: *buồng chuối*; of areca fruits: *buồng cau*), an ovary (*buồng trứng*)
bữa	Q	a meal (*bữa cơm*), a banquet (*bữa tiệc*), a day (*bữa sinh-nhật* 'birthday')
bức	P	a wall (*bức tường*), partition (*bức vách*), curtain (*bức màn*), painting (*bức tranh*), letter (*bức thư*), telegram (*bức điện-tín*)
canh	Q	watch (of the night)
cánh	P	a wing (of a door: *cánh cửa*), a sail (*cánh buồm*), a field (*cánh đồng*)
cành	Q	a branch, bough (of tree: *cành cây*), a spray (of flowers: *cành hoa*)
cặp	P	a pair (of eyes: *cặp mắt*), a couple (unmarried; *cặp trai gái*; married: *cặp vợ chồng*; of lovers: *cặp uyên ương*)
cây	P	a tree or plant (See 1.2), a plantlike object, such as a candle: *cây nến*.
cân	M	a catty—equivalent to 1.33 pounds—or a kilogram (of meat: *cân thịt*). See *lạng*, *tạ*
chai	T	a bottleful (of wine: *chai rượu*)
chậu	T	a washing-bowlful (of water: *chậu nước*)
chén	T	a cup(ful) (of tea: *chén trè*)
chiếc	P	one of a pair (of chopsticks: *chiếc đũa*)

[17] Cf. 4.2.2.4.

chiếc	P	a vehicle (*chiếc xe*), a boat (*chiếc thuyền*), a ship (*chiếc tầu*), an airplane (*chiếc máy bay*), a bridge (*chiếc cầu*)
chòm	P	a cluster (of trees: *chòm cây*)
chỏm	P	a mountain (*chỏm núi*)
chục	M	a group of ten (tangerines: *chục quít*)
chum	T	an earthenware jarful (of water: *chum nước*)
chùm	P	a bunch (of grapes: *chùm nho*; of keys: *chùm chìa khóa*)
chuyến	P	a trip, a round, a cartage (*chuyến đi*, *chuyến xe*), a train (*chuyến xe lửa*)
chương	P	a chapter (of a book: *chương sách*)
con	P	See 5.1
cỗ	P	a carriage—probably complete with horses (*cỗ xe*); a coffin (*cỗ áo quan*)
cốc	T	a glassful (of champagne: *cốc rượu sâm banh*; of wine: *cốc rượu*; of water: *cốc nước*)
cơn	P/V	a fit (of anger: *cơn giận*); a squall (of rain: *cơn mưa*), etc.
củ	P	a bulb (of onion: *củ hành tây*); a tuber (of manioc: *củ sắn*)
cuốc	P	a trip, a ride (*cuốc xe*)
cuộc	P	a party, a game (of chess: *cuộc cờ*); a meeting (*cuộc mít-tinh*); a contest (*cuộc thi*), an election (*cuộc tuyểncử*), a revolution (*cuộc cách-mạng*), a life (*cuộc đời*), etc.
cuốn	P	a fascicle, a book, a volume (*cuốn sách*)
cuộn	P	a roll (of paper: *cuộn giấy*; of cloth: *cuộn vải*; of film: *cuộn phim*)
đám	Q	a crowd (of people: *đám người*) gathering for certain rituals; a cloud (*đám mây*; an agglomeration (of trees: *đám cây*; of houses: *đám nhà*, etc.)
đàn	Q	a herd (of buffaloes: *đàn trâu*, etc.) a flock (of birds: *đàn chim*, etc.), a school (of fish: *đàn cá*), a crowd (of people: *đàn người*, etc.)
đạo	P	a law (*đạo luật*), an order (*đạo nghịđịnh*), a decree (*đạo sắc-lệnh*), etc.
đấu	M	a peck—316 cubic inches (of corn: *đấu ngô*)
đĩa	T	a plateful (of meat: *đĩa thịt*)
điếu	P	a cigarette (*điếu thuốc lá*), a pipe (of tobacco: *điếu thuốc lào*; of opium: *điếu thuốc phiện*)
đỉnh	P	a summit (of a mountain: *đỉnh núi*)

đóa	P	a flower (*đóa hoa*)
đoạn	P	a section, paragraph, passage (of a book)
đôi	P	a pair (of shoes: *đôi giầy*; of chopsticks: *đôi đũa*), a couple (husband and wife: *đôi vợ chồng*)
đội	Q	a company, a detachment (of soldiers: *đội lính*)
đống	Q	a heap (of refuse: *đống rác*), a pile (of sand: *đống cát*), a stack (of straw: *đống rơm*)
đồng	P	a coin (for a half-penny: *đồng chinh*; for a penny: *đồng xu*; for a dime: *đồng hào*; for 20 cents: *đồng ván*; for a piastre: *đồng bạc*); a piastre's worth of
gang	M	a span—space between the end of the thumb and the end of the middle finger when extended (*gang tay*)
gánh	M/P	a pole-load—quantity that can be carried in two containers (baskets, etc.) suspended from both ends of a carrying pole, (*gánh lúa*: two basketfuls of paddy); a troupe (of perambulating actors and actresses: *gánh hát*; of circus people: *gánh xiếc*)
góc	M	a corner, a piece (of a cake: *góc bánh*)
gói	T	a package (given as a gift: *gói quà*), a bundle (of clothes : *gói quần áo*), a pack (of cigarettes: *gói thuốc lá*)
gian	P	a room (*gian buồng/phòng*)
giấc	P/V	a sleep, a nap (*giấc ngủ*), a dream (*giấc mơ/mộng*)
giẫy	P	a row (of houses: *giẫy nhà*; of trees: *giẫy cây*; etc.)
giọt	P/M	a drop (of blood: *giọt máu*; of water: *giọt nước*; etc.)
hãng	Q	a business firm (*hãng buôn*)
hạt	P	a grain, a seed (*hạt giống*), a drop (of rain: *hạt mưa*; etc.)
hòm	T	a boxful; trunkful (of clothes: *hòm quần áo*). Cf. *rương*
hòn	P	a stone (*hòn đá*), a stone-like object (bullet: *hòn đạn*), an island (*hòn đảo*)
hồi	P/V	a chapter (of a novel: *hồi tiểu-thuyết*), an act (of a play: *hồi kịch*); a time, round
hộp	T	a boxful, a carton (of candies: *hộp kẹo*; etc.)
hụm	M	a gulp, a drink (of water: *hụm nước*; etc.)
ít	M	a small quantity (of salt: *ít muối*), a small number (of tools: *ít đồ dùng*; etc.)
khẩu	P	a mouthful (of sugar cane: *khẩu mía*)
khoảnh	M	an area equivalent to 100 *mẫu* '*mow*', that is 360,000 square meters

khúc	P	a portion, section (of a fish: *khúc cá*; of a tree trunk: *khúc gỗ*); a section (of a river: *khúc sông*; of a road: *khúc đường*); a song or a piece of music (*khúc nhạc*)
ki-lô	M	a kilogram (of fat: *ki-lômỡ*; etc.)
kỳ	Q	a session (of a meeting: *kỳ hội-đồng*); an examination (*kỳ thi*)
lá	P	a leaf (of banana: *lá chuối*); a card (in games: *lá bài*); a sail (*lá buồm*), a flag (*lá cờ*), a lung (*lá phổi*); etc.
lạng	M	a tael—equivalent to 37.8 grams or one sixteenth of a *cân* or a hectogram (of gold: *lạng vàng*; etc.) See *cân*, *lạ*.
lát	Q	See *nhát*
liễn	T	a porcelain jarful (of rice: *liễn cơm*; etc.)
lít	M	a liter (of kerosene: *lít dầu lây*, etc.)
loài	Q	a kind, sort (of weapons: *loại võ-khí*)
lời	Q	a word, an utterance, a statement (*lời tuyên bố*)
luồng	P	a gust, draft (of wind: *luồng gió*), a sound wave (*luồng sóng điện*); a current (of electricity: *luồng điện*)
lứa	Q	a hatch, a brood (of chickens: *lứa gà*; etc)
lượng	M	See *lạng*
ly	M	a millimeter
màn	P	scene (of a play: *màn kịch*)
mảnh	M	a piece (of wood: *mảnh gỗ*; etc.), a small field (*mảnh ruộng*), a small garden (*mảnh vườn*), a moon—not full (*mảnh giăng*)
mẩu	M	a piece (of bread: *mẩu bánh mì*, etc.)
mẫu	M	a mow—equivalent to 3600 square meters—or a hectare (of ricefield: *mẫu ruộng*; etc.) Cf. *sào*
miếng	M	a mouthful; a slice, a piece (of meat: *miếng thịt*, etc.)
miếng	P	a trick (of the art of fighting: *miếng võ*)
món	P	a sum (of money: *món tiền*), a debt (*món nợ*), a dish or course of food (*món ăn*), an item (of merchandise: *món hàng*), a subject (of study: *môn học*)
mối	P	a sentiment or feeling (of love: *mối tình*), a relationship (*mối quan-hệ*), a cause for (hope: *mối hi-vọng*; worry: *mối lo-ngại*), a danger (*mối nguy-hiểm*), a threat (*mối đe-dọa*), etc.
môn	Q	a subject, a field (of study: *môn học*); a specialty, an art

mớ	P	a tray (of roasted sticky rice: *mớ cốm*), a layer (of clothes: *mớ áo*), a bundle, a mass (of materials: *mớ tài-liệu*)
múi	P	a section (of grapefruit: *múi bưởi*; of tangerine: *múi quít*; of orange: *múi cam*; of jackfruit: *múi mít*; of mangosteen: *múi bứa*; etc.)
mũi	P	an arrow (*mũi tên*), a lance (*mũi giáo*)
muỗng	T	a spoonful (of salt: *muỗng muối*)
nãi	P	a hand (of bananas: *nãi chuối*). Cf. *buồng*
nạn	Q	an accident; a disaster, catastrophe, danger, peril
nắm	T	a handful (of raw rice: *nắm gạo*), a fist (*nắm tay*—*nắm đấm* is more aggressive), a ball (of cooked rice: *nắm cơm*)
nét	P	a stroke (of pen, pencil, brush: *nét bút*), a facial feature (*nét mặt*)
nền	Q	an economic, political or cultural institution (culture: *nền văn hóa*; civilization: *nền văn minh*; freedom: *nền tự-do*; democracy: *nền dân-chủ*; independence: *nền độc-lập*; economy: *nền kinh-tế*; trade: *nền thương-mại*; industry: *nền kỹ-nghệ*, etc.)
ngón	P	a finger (*ngón tay*), a toe (*ngón chân*); a trick (to bluff: *ngón bịp*)
ngọn	P	a mountain (*ngọn núi*), a hill (*ngọn đồi*), a tree (*ngọn cây*), a lamp (*ngọn đèn*), a flag (on pole: *ngọn cờ*), etc.
ngôi	P	a grave (*ngôi mả/mộ*); a star or a movie star (*ngôi sao*); a house, building (*ngôi nhà*); a temple (Buddhist: *ngôi chùa*; Confucianist: *ngôi đền*; Taoist: *ngôi miếu*; etc.)
ngụm	M	See *hụm*
nhà	T	a house, building (inn: *nhà trọ*; etc.), a houseful (of paddy: *nhà thóc*)
nhát	V	a cut, a stab, a wound (with a knife: *nhát dao*); a stroke (with a hammer: *nhát búa*)
nóc	P	a house, building (*nóc nhà*)
nụ	Q	a bud (of tea: *nụ chè*)
nương	Q	a terrace field (of sweet potatoes: *nương khoai*, etc.) Cf. *ruộng*
ổ	Q	a nest (of crows: *ổ quạ*; etc.), a brood (of chickens: *ổ gà*; puppies: *ổ chó*; of mice: *ổ chuột*; etc.), a lock (*ổ khóa*), an engine (*ổ máy*), a loaf of bread (*ổ bánh mì*)

ôm	M	a quantity that can be held in both arms (armful of books: *ôm sách*)
phát	P/M	a shot (of firearm: *phát súng*); an injection (*phát tiêm*)
phân	M	a fen—one tenth of a *tấc*, which is equivalent to 1.41 inches—or a centimeter (of land: *phân đất*); a fen—one tenth of a *tiền* 'mace,' which is equivalent to 3.78 grams—or a gram (of gold: *phân vàng*). See *tấc* and *tiền*
phần	Q	a part, section
phong	P	a letter—contained in an envelope (*phong thư*), a small package (of rice cake: *phong bánh*)
quả	P	a fruit, or a round, spheric object (heart: *quả tim*; kidney: *quả thận*; earth: *quả đất*; mountain: *quả núi*; hill: *quả đồi*; pendulum: *quả lắc*; ball: *quả bóng/ban*; bomb: *quả bom*; etc.)
quan	M	a string of coins (*quan tiền*)
quẻ	M/V	a divination diagram, a horoscope (*quẻ bói*)
quyển		See *cuốn*
rá	T	a quantity that can be contained in a small bamboo basket called *rá* (basketful of raw rice: *rá gạo*), whose weave is finer than in the one called *rổ* or *thúng*.
rặng	Q	a row or line (of trees: *rặng cây*)
rổ	T	a quantity that can be contained in a bamboo basket called *rổ* (basketful of raw rice: *rổ gạo*) whose weave is broader than in the one called *rá*, but which is smaller than the one called *thúng*.
ruộng	T	a field—usually in a low spot (of corn: *ruộng ngô*). Cf. *nương*
rương	T	a trunkful (of clothes: *rương quần áo*). Cf. *hòm*
sào	M	one tenth of a *mẫu*—i.e. 360 square meters (of ricefield: *sào ruộng*; etc.). See *mẫu*
sô	P	a sum (of money: *sô tiền*). Cf. *món*
sợi	P	a hair (*sợi lông*), a thread (*sợi chỉ*), a string, cord (*sợi giây*), a rope (*sợi thừng*)
tá	M	a dozen (eggs: *tá trứng*)
tạ	M	a picul—equivalent to 100 catties (i.e. 133.33 pounds)—or a metric ton (of rice: *tạ gạo*). See *cân*.
tách	T	a cupful (of tea: *tách trà*). Cf. *chén*.
tấc	M	a Vietnamese inch—actually 1.41 inches—or a decimeter (of land: *tấc đất*). See *phân*.

tấm	P	a bolt (of cloth: *tấm vải*, etc.); a board, plank (*tấm ván*); a hide (*tấm da*), a photograph (*tấm ảnh/hình*) a mirror or an example (*tấm gương*), a heart (*tấm lòng*), etc.
tấn	M	a ton (of coal: *tấn than*)
tấn	P	a play, a drama (*tấn tuồng, tấn kịch*)
tầng	P	a story (of a building: *tầng gác*)
tia	Q	a ray, beam (of light: *tia sáng*), a jet (of water: *tia nước*)
tiền	M	a mace—one tenth of a lạng 'tael', which is equivalent to 37.8 grams—or a decagram (of gold: *tiền vàng*). See *lạng, phân.*
tòa	P	a building (big house: *tòa nhà*; temple: *tòa miếu*) Cf. *ngôi, nóc*
tô	T	a bowlful (of noodle soup: *tô mì*)
tờ	P	a sheet (of paper: *tờ giấy*; etc.), a newspaper (*tờ báo*), etc.
thanh	P	a thin stick (of wood: *thanh củi*), a thin piece of material (a sword: *thanh kiếm*; a sabre: *thanh gươm*; a scimitar: *thanh đao*; a lance: *thanh giáo*; etc.)
thìa	T	a spoonful (of sugar: *thìa đường*, etc.) Cf. *muỗng*
thiên	P	(a chapter of) a novel (*thiên tiểu-thuyết*)
thúng	T	a market basketful (of eggplant : *thúng cà*)
thùng	T	a bucket (of water: *thùng nước*), a can (of trash: *thùng rác*), a barrel (of wine: *thùng rượu*)
thứ	Q	a kind, sort, category (a language: *thứ tiếng*)
thửa	Q	a plot (of land: *thửa đất*; of ricefield: *thửa ruộng*).
thước	T	a Vietnamese foot—equivalent to 14.1 inches—or a meter (of silk: *thước lụa*, etc.)
trái	P	See *quả*
trận	Q/V	an outburst (of rain: *trận mưa*), a contest, struggle, battle (war: *trận chiến-tranh*)
vại	T	an earthenware jarful (of salted vegetables: *vại dưa*)
viên	P	a pill, a tablet (*viên thuốc*), a bullet (*viên đạn*), a brick (*viên gạch*), a tile (*viên ngói*). See *hòn.*
vở	P	a play: *vở tuồng, vở kịch*
vụ	P	an unfortunate happening (flood: *vụ lụt*; fire: *vụ cháy*; robbery: *vụ trộm*; an assassination: *vụ ám-sát*; bomb raid: *vụ ném bom*; lawsuit: *vụ kiện*; etc.)
vũng	T	a holeful (of water: *vũng nước*)

vườn	T	a garden (ful) (of vegetables: *vườn rau*; of flowers: *vườn hoa*; of guavas: *vườn ổi*)
vuông	P	a square (of cloth: *vuông vải*)

7. The use of classifiers is a feature which Vietnamese shares with several languages of Far and South Eastern Asia, for instance, Chinese,[18] Japanese,[19] Korean,[20] Thai,[21] Burmese,[22] and Malay.[23] As Mary R. Haas says, "a comparative, or better, contrastive, exposition of the variations in usage among the several languages would make an interesting and valuable study."[24]

[18] Charles F. Hockett, *Spoken Chinese* (New York: Henry Holt Co., 1944). pp. 24-25, or Yuen Ren Chao, *Mandarin Primer* (Cambridge, Massachusetts: Harvard University, 1948), pp. 45-46. Also Edward H. Schafer, Jr., "Noun Classifiers in Classical Chinese," *Language*, 24 (1948), pp. 408-413.

[19] Bernard Bloch and Eleanor Harz Jorden, *Spoken Japanese* (New York: Henry Holt Co., 1946), p. 162.

[20] Fred Lukoff, *Spoken Korean* (New York: Henry Holt Co., 1945), pp. 219-220.

[21] Mary R. Haas and Heng R. Subhanka, *Spoken Thai* (New York: Henry Holt Co., 1945), pp. 59-61. Also Mary R. Haas, "The Use of Numeral Classifiers in Thai," *Language*, 18 (1942), pp. 201-205.

[22] William S. Cornyn, *Spoken Burmese* (New York: Henry Holt Co., 1945), pp. 144-146, or William S. Cornyn, *Outline of Burmese Grammar* (Language Dissertation No. 38, Baltimore: Linguistic Society of America, 1944), pp. 26-28. Also Mary R. Haas, "The Use of Numeral Classifiers in Burmese," *Semitic and Oriental Studies*, 11 (1951), pp. 191-200.

[23] Isidore Dyen, *Spoken Malay* (New York: Henry Holt Co., 1945), pp. 114-115.

[24] Mary R. Haas, *op. cit.*, p. 191.

Faculty of Letters.
University of Saigon.

DEUX SYSTÈMES DE VOCABULAIRE PARALLÈLE A MADAGASCAR

JOSEPH VERGUIN

Nous présentons ici quelques résultats obtenus au cours d'une enquête menée dans l'Andruy, Sud de Madagascar, en 1955.

Le dialecte tandruy possède un système de vocabulaire double, que nous appelons « vocabulaire parallèle ». Sur les cent trente mots servant à l'enquête, trente-cinq entrent dans ce système. Ce nombre n'est nullement limitatif. En effet nos informateurs nous ont indiqué spontanément d'autres exemples qui ne figuraient pas dans notre questionnaire. Cela ne signifie pas non plus que le système s'étende à la majeure partie du vocabulaire. Il semble bien, d'après le relevé que nous avons pu faire, qu'il intéresse surtout ce qui touche directement l'individu: parties du corps, fonctions naturelles, objets les plus familiers.

Nous sommes ici en présence d'un phénomène social. A l'heure actuelle il est difficile d'obtenir des informateurs une explication claire du fait en cause. Lorsque nous avons questionné hommes ou femmes à ce sujet ils ne semblaient plus très informés sur le sens profond d'une opposition pourtant encore vivante pour eux. D'une façon générale ils conviennent que le terme de la première série s'applique aux jeunes ; et le terme de la seconde série revient de droit aux « aînés » (proprement aux « undati be »). Tel est en effet, semble-t-il, l'état actuel des choses. Le locuteur lorsqu'il s'adresse à une personne d'une génération aînée, ou lorsqu'il parle d'elle, emploie le terme de la seconde série, que l'on est tenté de définir ainsi comme un terme de respect. Par contre s'adressant à un contemporain ou à un cadet, il emploie le terme de la première série.

C'est ainsi que la phrase : « Tu *manges* de la viande » se traduira par « *Mikama* hie hena » à l'adresse d'un aîné, et par « *huma* hie hena » dans le cas contraire. De même « *Mikama* ize hena » (il mange de la viande) parlant d'un aîné, par opposition à « *Huma*

ize hena ». Ce parallélisme nous semble s'expliquer par la structure traditionnelle de la société. La grande tribu tandruy est constituée de nombreux clans groupant des familles d'une même origine. Autrefois chaque clan obéissait à un chef ou « Ruandria ». Existait également un chef religieux le « Mpisuru » et, en temps utile, on choisissait pour son courage un chef de guerre, le « Talē ». Au sein de la famille le père demeure le maître absolu. D'autre part, d'après Defoort, la société était divisée autrefois en trois classes : les « Andria », les « Vuhitse » et les « Undevu » ; ce qui correspond approximativement aux nobles, bourgeois et esclaves. Cette hiérarchie sociale avait son reflet dans la langue où les termes spéciaux étaient réservés aux chefs et, si ce qu'avance Defoort est exact, le système que nous observons aujourd'hui devait alors jouer entre les classes.

C'est ce qui apparaît encore chez les Masikuru[1], dans le Bas-Mangoky, où les termes de la seconde série s'appliquaient strictement aux roitelets locaux. Cette série, utilisée parfois encore à l'égard des descendants des derniers rois, tombe néanmoins dans l'oubli. Un de nos informateurs, fils d'un ancien roi, et âgé d'une quarantaine d'années a dû, pour certains termes, faire appel à sa mère présente, tandis qu'un vieillard, autre informateur, avait gardé bonne mémoire de son système. Dans certains villages la seconde série était réduite à un ou deux termes.

Mais la structure sociale traditionnelle tandruy s'est trouvée atteinte par la pénétration française. Une organisation administrative nouvelle s'est installée. Les termes consacrés aux anciens chefs ne sauraient être réservés avec la même valeur aux chefs de village, de quartier ou de cantons actuels.

Il semble bien que petit à petit les termes de respect aient dû s'étendre à ceux vers qui allait automatiquement le respect de tous : les vieillards, les aînés.

Nous présentons ici, en juxtaposition, les systèmes parallèles tandruy et masikuru. On notera la quasi identité des deux systèmes pour la série 1, alors que les similitudes sont relativement rares dans la deuxième. Il est toutefois intéressant de remarquer la fréquence du préfixe *fi-* dans les termes de la seconde série, dont certains sont dérivés au moyen de ce suffixe de la première série.

[1] Il est généralement admis que les Tandruy sont les descendants des Sakalava émigrés de la côte Ouest vers le Sud de la Grande Ile et que les Masikuru du Bas-Mangoky sont eux-mêmes de la grande tribu Sakalava.

	TANDRUY		MASIKURU	
	1re série (jeunes)	*2e série* (aînés)	*1re série* (commun)	*2e série* (roi)
tête	*luha*	*ambuñe, añabu*	*luha*	*kabesu*
cheveux	*vulu*	*marui*	*vulu*	*marumana*
yeux	*masu*	*fihainu*	*masu*	*fidzeri, fizaha*
nez	*uru*	*fihantsuna*	*uru*	*fañimbu*
bouche	*vava*	*falie*	*vava*	*tsihuvi*
dents	*nife*	*fihitsike*	*hihi*	*salake*
oreilles	*sufi*	*ravembia, ravimbia fivimbi*	*sufi*	*fitilu, fiharaña (fidzina)*
langue	*lela*		*lela*	*fañamitsi*
cou	*vuzu*	*fititia lanteahue*	*atuke*	*fituliha*
poitrine	*tratra*	*araña falurañа fihufa*	*tratra*	*mandzehadzeha*
sein	*nunu*	*fatrua*	*nunu,*	*tsepuke*
bras	*a|taña b|sira*	*a|fita (néant)*		
main	*taña*	*fitā*	*taña*	*fandrambe*
doigts	*rambulaña*	*rambumpitā*		
ventre	*truke*	*fisafua*	*kibu*	*truke*
nombril	*fuelse*		*kubu*	*tsimañadrue*
pied	*tumbuke*	*fandia*	*tumbuke*	*fañitsaka*
maison	*traña*	*andzumba*	*traña*	*zumba*
tombeau	*lulu*		*lulu*	*aranka*
époux	*vali*	*lanadzumba*		
homme	*lahilahi*	*ruandria rangahe andruanavi*		
femme	*ampela*	*ruakemba rakemba*		
manger	*huma*	*mikama*	*huma mihina(na)*	*mikama midzabu*
boire	*minundranu*	*mikama ranu*	*minundranu*	*miluhuke*
cracher	*mandrura*	*fañare milsurike mañare mizake*		
dormir	*miruru*	*mirulse*	*miruru*	*mirulse*
uriner	*mamani*	*mañariranu*		
avoir faim	*kerē*	*musare*		
avoir soif	*marandranu*	*laliñere*		
mourir	*male*	*vilasi*		
parler	*mivula, mireake*	*misauntsi, mitsara*		
rire	*humehe*	*miankahake*		
pleurer	*lumañe*	*miaruvitse*		
venir	*āvi*	*mitunga*		

	TANDRUY		MASIKURU	
	1re série (jeunes)	*2e série* (aînés)	*1re série* (commun)	*2e série* (roi)
partir	*mandeha*	*mañavelu*		
marcher	*mandeha*		*mandeha*	*milandava*
s'asseoir	*milubuke*	*miambesalse*		
se marier	*manambali*	*manalanandzumba*		

Centre National de la Recherche Scientifique, Paris.

REVIEWS

Diego Catalán Menéndez-Pidal, *La escuela lingüística española y su concepción del lenguaje.* Biblioteca Romanica Hispanica. Madrid : Editorial Gredos, 1955.

Si le livre de Diego Catalán ne réussit pas à convaincre le lecteur de l'existence véritable d'une « école espagnole », il présente cependant avec à propos et de façon systématique la doctrine de Ménendez-Pidal et d'Amado Alonso. L'école espagnole, puisque l'expression est commode, illustre une conception spiritualiste du langage, par opposition à la conception historiciste des néo-grammairiens et à la conception sociologique de l'école française. L'activité de l'esprit s'impose comme thème essentiel de la linguistique. Cette prise de position amène à concevoir la langue non pas comme *ergon*, mais comme *energeia*, et par conséquent à orienter les recherches dans le sens diachronique et à ne pas séparer philologie de linguistique, puisque la langue littéraire ne diffère pas en son essence de la langue parlée.

La langue est un système limité de signes arbitraires, cependant les sujets parlants, loin d'être conscients de l'arbitraire des signes qu'ils emploient, cherchent constamment à les motiver ; bien que ce sentiment de motivation ne soit à proprement parler qu'une illusion, il faut compter avec lui en linguistique, en tant qu'activité de l'esprit. Le système n'est d'ailleurs pas conçu de façon rigide : la dichotomie saussurienne langue-parole fait place à la triple distinction entre langue-parole-norme. Si on se contente d'une linguistique immanente du type préconisé par Saussure, en écartant résolument toute considération diachronique, on se trouve dans l'impossibilité d'expliquer l'existence de la norme, « cette entité supra-individuelle qui s'impose à tous les membres d'une même communauté linguistique ». Toute étude synchronique repose sur une abstraction qui essaie de réduire à un système de signes unique ce qui est, dans chaque sujet parlant, une lutte entre plusieurs systèmes simultanés et souvent incompatibles.

Les linguistes espagnols font donc de l'étude des changements linguistiques l'objet de leurs recherches et prennent pour guides les processus de création et de transmission qu'ils observent dans le domaine de la poésie traditionnelle. Le conflit entre l'individuel et le social, la nécessité de rendre social, à des fins communicatives, ce qui est individuel, fait qu'il est dans l'essence même de la langue de changer ; mais de même que, pour un parler donné, les variantes sont le résultat d'un acte créateur de la part du chanteur, de même le changement linguistique a son origine dans l'initiative d'un locuteur isolé que le besoin d'expressivité pousse à innover. Cependant, de même que certaines variantes du même parler restent localisées tandis que d'autres se répandent, de même il n'y a qu'un très petit nombre d'initiatives linguistiques individuelles qui soient acceptées par la communauté. Il faut donc étudier à part ce qui permet l'initiative individuelle et ce qui favorise sa généralisation.

Diego Catalán mentionne au passage l'existence de facteurs de changement inhérents au système (économie, rendement fonctionnel, etc.), mais on voit bien que ce qui intéresse les linguistes espagnols, c'est bien plutôt l'étude des causes externes, qu'elles soient d'ordre psychologique (tendance à l'expressivité, etc.) ou culturel (contacts entre dialectes ou entre langues, etc.). Le processus du changement est vu comme une lutte séculaire entre des forces conservatrices (*cultismo*, risques d'homonymie, etc.) et des forces néologiques, dont l'étude se ramène à l'examen des courants culturels qui ont fait prévaloir les unes sur les autres.

Le mérite des linguistes espagnols en tout ceci est de s'élever contre le principe du caractère aveugle des changements linguistiques, contre la dichotomie changements réguliers-changements analogiques, et de montrer que le changement linguistique n'est ni une dégradation, ni quelque chose de brutal, mais un très lent processus qui fait que dans un état de langue donné coexistent des vestiges d'états passés et des embryons d'états futurs.

Les études diachroniques s'orientent maintenant vers l'analyse des causes internes des changements linguistiques, mais il faut remarquer que l'extrême ouverture d'esprit avec laquelle les linguistes espagnols ont ménagé une place à ce genre de travaux, tout en s'intéressant eux-mêmes à des problèmes totalement différents, préserve leurs recherches du risque de perdre leur valeur, une fois confrontées avec les tendances nouvelles. C'est cette même ouverture d'esprit qui caractérise le petit livre dont Diego Catalán a voulu faire une introduction à la linguistique

générale. De ce dernier point de vue, il est fâcheux que le point de départ aussi résolument spiritualiste lui fasse pratiquement nier la valeur des études synchroniques et négliger l'étude des systèmes en eux-mêmes, à des fins diachroniques aussi bien que synchroniques.

Sceaux. GENEVIÈVE CORRÉARD.

W. FUCKS, *Mathematische analyse von Sprachelemente, Sprachstil und Sprache.* 110 pp., Köln : Westdeutscher Verlag, 1955.

Many independent statistical investigations of language blossom into books, these days. The variety and lack of overlapping of these books is quite striking. The book by Fucks contains some bits of theory, of which all we wish to say is that we strongly disapprove of them, and do not see their point. But it contains mostly an enormous collection of 70 graphs or tables of miscellaneous data. As a collector of data, Fucks belongs to the same redoubtable tradition as G. K. Zipf: he seems unable to stand a theoretical coefficient that has not been calculated in an enormous number of cases. Here he has listed data concerning: the convergence of statistical characteristics to the mean values around which they are supposed to be scattered; the distribution of the frequencies of various numbers of syllables per word, mostly in German and Latin texts; the distribution of the frequencies of various meters of poetic rythm, also in German and Latin; the transition frequencies between various kinds of elements, that is: the frequencies with which certain elements occur in some places when it is known that certain others occur in previous places; various indices to describe these transition frequencies, which are rather bulky to write down; comparisons between several languages from such viewpoints as the relationship between the entropy per syllable and the average number of syllables per word; and the diachronic study of German and Latin from the viewpoint of the evolution of the entropy and the average number of syllables per word, during the development of these languages; the trend is not striking.

The aim of Fucks, as that of his predecessors, G. Udny Yule and P. Guiraud, is mostly the analysis of style, and for that he tries to find numerical characteristics of speech, which differ as strikingly as possible between different authors. The same requirement is to be faced in studies tending to make it easier to

recognize the authorship of texts. There is little doubt about the practical importance of such studies, and the only way to handle the problem today, is to try the greatest possible number of numerical characteristics for their efficiency as discriminators. But the reviewer cannot help feeling that it is very much more needed to first study the numerical features that are *common* to as many texts as possible. Of course, one never knows in advance whether a feature will be common or not: our feeling should be understood as an encouragement to students to investigate preferably those features of language, that appear at first glance to be common to many texts or authors.

We should like to end with a warning — which does not apply to Fucks — concerning the use of information theory in linguistics. There is no point whatsoever in computing entropies of various linguistic elements, when one does not need them for some further aim, unless one wants to do service to further investigators. The most useful in information theory is not, by any means, the definition of information-entropy, but a structural and formal way of thinking about language, which only a greater familiarity with the theory can bring and which carries great promise of further theoretical ideas about what to count: such ideas are what is most lacking now in statistical linguistics.

University of Geneva. B. MANDELBROT.

MARGARET SCHLAUCH, *The Gift of Language*, IX+342 pp. New-York, Dover, 1955.

Madame M. Schlauch, actuellement professeur à l'Université de Varsovie, résidait autrefois en Amérique, où elle a publié en 1942 un excellent livre de vulgarisation scientifique « The Gift of Tongues ». Il semble que nous ayons aujourd'hui la simple réedition de cet ouvrage, car, dans la bibliographie, nous ne trouvons aucune référence postérieure à 1942. Il n'est donc pas trop étonnant que cet ouvrage ne fasse pas allusion aux thèmes de la linguistique structurale actuelle.

Après une brève introduction sur l'origine et la fondation du langage (p. 1-18), la linguistique synchronique est traitée assez largement, en même temps que la classification des langues (p. 19-169). La linguistique diachronique est exposée plus brièvement sur la base de la phonétique historique de l'anglais et de l'histoire de cette langue (p. 170-226). Les derniers chapitres sont peut-être

plus originaux : l'un étant consacré à la création poétique, utilisation individuelle affective du langage ; l'autre à l'aspect social de celui-ci dans nos sociétés.

Malheureusement, comme nous l'avons dit, la bibliographie (p. 299-306) n'est pas à jour. Faut-il en conclure que cette réédition a été faite sans l'autorisation de l'auteur ? Il est à craindre qu'il en soit ainsi, si, comme on le pense, la propriété des œuvres n'est pas aussi efficacement protégée aux U. S. A. qu'en Europe.

André G. Haudricourt.

G. Devoto, *Dizionari di ieri e di domani*, 125 pp., Firenze: G. C. Sansoni, no date.

The pamphlet is the reprint of three articles, or rather, long reviews, published between 1935 and 1943: one of Panzini's *Dizionario moderno* (pp. 7-53); one of the (now dead) *Vocabolario dell'Accademia* (pp. 55-85); and one of several etymological dictionaries, such as Meyer-Lübke, Ernout-Meillet, Walde, Bloch-Wartburg, Gamillschegg *(sic.)*. This gives the author the opportunity to present a scheme of how dictionaries, both etymological and non-etymological, should be composed; which is, as we know, one of the most difficult tasks that can confront a linguist. It is needless to say that all of Devoto's counsels are pure gold. The comparison between Walde and Gamillscheg on the one hand, Ernout-Meillet and Bloch-Wartburg on the other, is definitely in favor of the second pair; especially because they contain more useful and more pertinent, if less, information. Comparing e.g. the word *arista* in Ernout-Meillet and Walde (p. 96), Devoto remarks (p. 98) that Ernout-Meillet « contiene cinque cose di cui il Walde tace: i derivati e i composti; i suffissi analoghi in altre parole; l'indicazione del valore 'tecnico'; l'autore piú antico presso cui si ritrova (*sic*), Varrone; la sorte ulteriore, panromanza (salvo la Romanía) ». I would like to add two more, and most important ones: the geographic distribution and the social stratification (for the latter, cf. e.g. words like *rēs* or *flāmen* or *serō*, *sēuī* in Ernout-Meillet and in Walde, including Walde-Hofmann).

A small critical remark: ital. *madre*, *padre* (p. 107) are, in Italian of Gallo-Italic, or rather, of Gallo-Roman origin and are bad examples of the preservation in Italian of the Indo-European heritage.

In conclusion: a splendid booklet, which both those who write and those who read dictionaries would be wise to consult.

G. BONFANTE.

FERDYNAND ANTKOWSKI, *La Chronologie de la monophtongaison des diphtongues dans les langues indo-européennes* (Poznańskie Towarzystwo Przyjaciół Nauk, Wydział Filologiczno-Filozoficzny, Prace Komisji Filologicznej, Tom XVI, Zeszyt 4), pp. 53, Poznań: 1956.

The merit of a theory is judged by its productivity, by the fruitfulness with which its principles can be applied to unanticipated problems. The theory can be freed of the suspicion of being an *ad hoc* explanation of a limited set of facts or of a group of similar problems when it can be lifted from its intended context and used successfully elsewhere. This is a case in point. Ferdynand Antkowski has set for himself the task of establishing, *a priori*, the order in which, in any given language, diphthongs may be expected to become monophthongs. He bases his conclusions on certain principles taken from the domain of diachronic phonemics; namely, André Martinet's suggestion that the asymmetrical arrangement of the vocal organs provides less room for the back vowels than for the front, Haudricourt and Juilland's reduction of this difference to mathematical terms, and the functional desirability of maintaining useful phonemic oppositions.

Before introducing these principles from diachronic phonemics, Antkowski establishes certain phonetic criteria. The likelihood of the monophthongization of a diphthong is related to the articulatory distance between its two members. Thus *ei* is more likely to be simplified than is *ai*, because *e* and *i* are closer together than *a* and *i*, so that the assimilation of the former two could be accomplished in less time. Similarly, *ou* is likely to be simplified more quickly than *au*. Since the 'mixed' diphtongs *oi* and *eu* are composed of a front and back element that are not easily measured in these terms, Antkowski leaves them out of consideration. Hence his theory concerns only the Indo-European diphtongs *ei*, *ai*, *au*, and *ou*. At this stage we may conclude that *ei* and *ou* are likely to be simplified before *ai* and *au*. However the asymmetry must also be taken into consideration. Since the distance from *a* to *u* is shorter than the distance from *a* to *i*, and

the distance from *o* to *u* is shorter than that from *e* to *i*, we might expect (and this expectation is worked out mathematically by Antkowski, on the basis of the figures of Haudricourt and Juilland) that *ou* would be simplified before *ei*, and *au* before *ai*, giving a theoretical order of *ou*, *ei*, *au*, *ai*. But there is yet another problem: The more limited space available for the phonemes of the back series may somehow act as a deterrent to the production of new long vowels by monophthongization; the same pressure in the back series that seems to have produced a number of cases of the shifting of *u* to *ü* with reduction of the number of back vowels may serve just as well to inhibit an increase in that number. On this basis, Antkowski assumes that the necessity of preserving distinctions in the back series will outweigh the factor of the actual distance between the phonetic elements. Therefore the theoretical order in which we may expect the diphthongs to be monophthongized is *ei*, *ou*, *ai*, *au*.

Furthermore, it seems to work. In the second half of the paper Antkowski analyzes the treatment of diphthongs in a number of Indo-European languages, incidentally supplying occasional provocative insights into the relationship between the products of this process and the organization of the vocalic system as a whole. His results may be summarized briefly as follows.

In Greek, *ei* and *ou* are simplified at roughly the same time, shortly after the beginning of the historical period. *ai* is somewhat later, and *au* is never really monophthongized in the sense used here; its second element becomes a spirant.

In Latin, *ei* and *ou* at roughly the same time, *ai* somewhat later, *au* much later—in some dialects not yet.

In Indo-Iranian the shift of *e* and *o* to *a* brings with it the merger of *ei* and *oi* with *ai* and of *eu* and *ou* with *au*. *ai* and *au* subsequently become *e* and *o* respectively, there being no evidence as to the relative chronology of these shifts. Indo-Iranian, then, cannot be used to check the theory.

In Old High German, *ei* is clearly first, *ou* merges with *au* at an early period when *ŏ*>*ă* and is thereafter out of the picture, *ai*>*e* in some environments before the eighth century and in other environments creates a new *ei* during the ninth, *au* has shifted to *o* by the beginning of the ninth.

In Old Irish, the simplification of *ei* is pan-Celtic, *ou* is clearly second; as between *ai* and *au* the evidence is not very clear, but if anything *au* is last.

In Lithuanian too *ou* becomes *au* as *ŏ* becomes *ă*, leaving three diphthongs. Of these, *ei* and *ai* are both prehistoric, and the chronological difference is a matter for debate, but *au* survives intact into the historical period.

In Albanian and Slavic the relative chronology is not clear, but there is nothing to contradict the theory.

In Armenian, the shift of *ei* to *e* is pre-literary, as is the shift of *ou* to *oy* and *u*, but *ai* and *au* survive into classical Armenian. For the later developments, the evidence indicates that *au* was simplified in the twelfth century and *ai* somewhat earlier.

Insofar as Antkowski's examples go, then, the theory is affirmed in every instance, either explicitly or indirectly. If the theory continues to be affirmed through analysis of those later stages of the various branches of Indo-European (or other languages that present sufficient chronological depth) that went through another cycle of diphthongization and simplification, we shall be indebted to Antkowski for yet another insight into the subtle processes of phonological change.

WILLIAM DIVER.

HEINZ KRONASSER, *Vergleichende Laut- und Formenlehre des Hethitischen*. 292 pp. Heidelberg: Carl Winter, 1956.

The synchronic and comparative interpretation of Hittite evidence that Kronasser offers us here has amazingly little in common with the one Edgar Sturtevant presented in a long series of books and studies. For the late Yale comparativist, Hittite and the other Anatolian languages formed a genetic group that had, from a trunk common to them and traditional Indo-European, branched off so early that any comparison could only be achieved through a new proto-language, the common denominator of the old *Ursprache* and a new, Proto-Anatolian, one. For Kronasser, Hittite and its Anatolian neighbors are nothing but ordinary representatives of Proto-Indo-European. Where they are found to differ from the familiar Brugmannian reconstruction, they must have innovated, just as Greek innovated when it reduced its case pattern from eight to five units, or as Slavic did when it merged the old imperfects and aorists. This applies to phonology and morphology alike. For Sturtevant, both *ḫ* and the absence of a feminine gender were archaisms; for Kronasser, they are novel features resulting from contacts with non-IE neighbors or predecessors.

Younger American Indo-Europeanists, whose conception of Hittite and comparative linguistics have, so one-sidedly, been influenced by Sturtevant's teachings, should, by all means, read Kronasser's exposition so as to restore the balance. They will thereby get closer to one important aspect of the truth, namely the fact that present-day Hittite evidence is still so lacunar that most of what comparativists can do with it remains in the realm of hypothesis.

Kronasser's approach to IE comparison will, by many, be branded as hopelessly retarded. This reviewer would rather venture to characterize it as prejudiced in some respects only: his unconvincing plea against some novelties should not make us forget some of his analyses (that of *r*-forms, pp. 203-205, for instance) where a neat marshaling of the data leads to conclusions with which any Indo-Europeanist, be he of the most progressive brand, could readily agree.

But one chapter, at least, is quite disappointing. It is the 21-page section devoted to the laryngeal hypothesis. We have had lately quite a few expositions of the laryngeal theory by some of its able supporters, and even people as definitely committed to it as this writer could look forward, for a change, to a well-launched attack against the laryngeal edifice. But Kronasser, who, in a general way, is clearly hostile, is not quite willing unconditionally to reject the hypothesis ("Hier soll nicht die Laryngalhypothese als solche *unbedingt* abgelehnt werden", p. 78; the reviewer's italics). What he wants to show is that Hittite *ḫ(ḫ)* can not be interpreted as the reflex of some former phonemes whose existence however he is not ready to deny. Writing Hittite *ḫ* without the subscript crescent enables him to identify it with whatever is written *h* in the most varied languages, and leads him to explain it away as a hiatus breaker. All this is as far-fetched as can be. One fails to grasp why Kronasser, who has just written "Dass es untergegangene Phoneme gegeben hat... is wohl richtig", should feel compelled to build up the most improbable set of hypotheses just because he has, for some unknown reason, decided that the *ḫ*'s of *ḫanti*, *ḫastai*, *puḫḫar* can not be the reflexes of these phonemes. There is a risk that this piece, coming as it does early enough in the volume, might induce some readers to stop short and condemn the book as hopelessly biased, which would be a real pity.

ANDRÉ MARTINET.

HERMANN BERGER, *Zwei Probleme der millelindischen Laullehre, Münchener indologische Studien*, Heft 1. 96 pp. München: J. Kitzinger, 1955.

Ainsi que le souligne dans la préface le Professeur H. Hoffmann, cet ouvrage est le premier d'une série qui doit rassembler les travaux du séminaire d'indologie et d'iranistique de l'Université de Münich ; il renoue, par son objet, avec la tradition des grands maîtres E. Kuhn et W. Geiger ; l'auteur aborde en effet deux problèmes de phonétique moyen-indienne. A juste titre, il regrette dans son introduction l'insuffisance des études portant sur la phonétique du moyen-indien, insuffisance qui ne laisse pas d'offusquer non seulement la connaissance des langues néo-indiennes, mais encore l'appréciation même du sanscrit ; à trop considérer ce dernier comme un outil privilégié de la comparaison indo-européenne, on oublie souvent qu'il n'a été qu'une étape, particulièrement remarquable il est vrai, dans un long développement dont les phases suivantes méritent à certains égards une égale attention de la part des indianistes.

La première étude a pour objet le traitement en pāli du *ṛ*, représenté par *a*, *i*, *u*. Que la sonante-voyelle ancienne soit représentée en indien par une voyelle pure est déjà un phénomène remarquable ; mais en outre la répartition des timbres vocaliques paraissait due au seul hasard. J. Bloch (*Indo-aryen*, p. 36) en soulignait le caractère arbitraire en affirmant : « la coloration de la voyelle est imprévisible ». H. Berger semble vouloir relever ce défi.

Après un court chapitre comportant des remarques préliminaires sur la phonétique et qui constitue en partie un discours de la méthode, l'auteur essaye de justifier les traitements *a*, *i*, *u* par l'entourage phonétique (p. 21). Il continue ainsi les travaux de Johansson (1889) depuis lequel rien de véritablement neuf n'avait été apporté sur la question. Lorsque aucune action particulière ne s'exerce, *ṛ* est représenté par *a* (*vṛtra->vatra*, *vṛka->vaka*, p. 20) ; s'il y a un *i* dans la syllabe suivante (*ṛṣi->isi*), si *ṛ* suit une palatale *(śṛṅga->siṅga)* ou une gutturale palatalisée (p. 30-31), il est passé à *i ;* il devient *u* si la syllabe suivante contient un *u (pṛthu->puthu)* et après consonne labiale (p. 51 sqq.).

Malheureusement cette distribution si satisfaisante ne va pas sans maintes exceptions ; elle aurait été troublée par un jeu complexe — trop complexe pour être convaincant — d'actions diverses, de conditions particulières qu'il appartient aux spécialistes de

discuter. Le critère de l'expressivité (p. 35-37) paraît singulièrement hasardeux. De même, la justification du traitement par *i* du *ṛ* initial (sauf devant *u*) demanderait plus ample commentaire. On peut être en effet surpris, au premier abord, de lire « die Ruhestellung der Sprachorgane war im m.i. wie in vielen anderen Sprachen die i-Stellung ». Du *Taittirīya-Prātiśākhya* (II, 20-21) à Sievers (*Gr. d. Lautphysiologie*, p. 15) c'est la position de *a*, qui est décrite comme position « neutre », position « de repos », au moins en ce qui concerne la langue.

Dans la deuxième partie, l'auteur étudie le sort en m.i. du groupe *kṣ* dont le traitement n'est nulle part uniforme, et dont on a, dès les textes védiques, des exemples de notation par *cch* (AS *ṛcchárā* = VS *ṛkṣálā-*, AS mss *parcchit* = *parikṣít-*). Reprenant la doctrine classique (J. Bloch, *Les inscriptions d'Asoka*, collection Sénart, p. 56) H. Berger reconnaît que le traitement *cch* est occidental, le traitement *kkh* oriental. Mais les exceptions sont fréquentes. Aussi H. Berger tente-t-il d'expliquer la répartition des deux traitements non seulement par des raisons géographiques, mais aussi par la phonétique : ainsi la palatalisation de pāli : *kacchā* (sk. *kakṣā*) serait due à l'action dissimilante de la gutturale initiale, le *ch* de pāli *chamā* « sur la terre » au fait que le terme remonte non à *kṣamā́* mais à la forme védique *kṣmā́*.

Seul un spécialiste pourrait dire si l'auteur n'a pas eu dans ses démonstrations plus d'ingéniosité que de bonheur ; mais tel qu'il est, l'ouvrage de H. Berger constitue au moins une base pour les discussions et les progrès futurs.

Université de Lyon. J. Manessy-Guitton.

Vittoria Moltoni, *Gli influssi dell'osco sulle iscrizioni latine della Regio I* (Istituto lombardo di scienze e lettere, Estratto dai *Rendiconti*, Classe di lettere—Vol. LXXXVII—1954), 39 pp. Milan, 1954.

The traditional interest of Italian scholars in the interaction of Latin and the Italic dialects continues to produce new data and rewarding, if controversial, theories. In the present work we find an analysis of all the errors made in the Latin inscriptions of Augustus' Regio I, the territory extending from just south of Rome to about Salerno. Since this is the area (particularly the Campanian part of it) from which come some of the most important Oscan inscriptions, Oscan influence is to be expected, and the

evidence presented confirms that supposition. The author's chief interest seems to be the ultimate influence of Oscan on the southern Italian dialects of today, but her material will also be worth examination by linguists in the more general field of bilingualism and language contact, and it may be worth while, in this review, to reconsider some of the data from that point of view.

Many of the mistakes in these inscriptions are just what one might expect from a person speaking (or chiselling) Latin with Oscan phonetics and the background of the Oscan phonemic system. There is considerable confusion in the handling of the higher front vowels, an area in which the Oscan and Latin patterns do not correspond at all. As Ralph Ward showed some years ago in an article in *Language* XIX (1943), 152-154 (an article that seems to have been generally overlooked by Italian linguists) Oscan had three relevant degrees of aperture, with a long and short vowel for each, where Latin had only two. In Oscan inscriptions in the native alphabet there is a special letter, transcribed *i*, to indicate the middle of these three degrees of aperture, but in inscriptions in the Latin alphabet the Oscan *i* is represented variously and inconsistently by Latin *i* and *e*. From the present study it becomes apparent that the Oscan-speakers were having the same trouble with Latin phonemes that they had had with Latin letters; Latin phonemes that fell within the phonetic range of Oscan *i* were likely to be rendered by either *i* or *e*. Thus Latin *ē*, which was phonetically closer than *ĕ*, is frequently written *i*, and Latin *ĭ*, which was phonetically more open than *ī*, appears as *e*: *ficil* for *fēcil* and *ameltil* for *amĭltit*.

Other errors seem to be the direct result of submitting Latin phonemes to Oscan rules of distribution. In Oscan, *e* seems never to stand as the first vowel in hiatus, and that such Latin combinations were pronounced in the Oscan manner is indicated by a number of words in which Latin *ea* is given as *ia*, such as *iamus* for *eamus*. Certain Latin consonant clusters that do not appear in Oscan receive similar reinterpretation: *uerecundus* becomes *uerecunnus*. A more complicated case is seen in the treatment of the cluster *ks*. Here we see not only the simplification of the cluster (*cacatris* for *cacatrix*), but also cases of hyper-correction that indicate that the Oscans of this area were aware that *x* was pronounced in some words, but that they did not know in which. Hence *milex* for *miles* and evan *xanto* for *santo*. Because of a similar uncertainty about anaptyctic vowels there is a considerable inconsistency in the presence or absence of vowels

in the neighborhood of liquids: both *specla* for normal Latin *specula* and *tempuli* for normal Latin *templi*. Moltoni groups examples of the former type under the heading of syncope. There certainly was plenty of syncope in the earlier history of Oscan, but it seems unlikely that it was still an active force in the Oscan distributional pattern at a time when the opposite development of anaptyxis was taking place. It would seem better to treat these apparent examples of syncope as a further indication of hypercorrection, particularly since every case of syncope cited by Moltoni takes place next to a liquid.

The treatment of the Latin clusters *kl* and *ps* is a problem to which there does not seem to be a clear solution. These clusters had been eliminated from Oscan during the prehistoric period; in the Latin inscriptions under consideration there are such errors as *uitoria* for *uictoria* and *isse* for *ipse*. The puzzle arises from the fact that these clusters actually did exist in Oscan (produced secondarily through syncope), so that there should have been no difficulty in pronouncing them. But then why were they simplified? There seems no reason to accept Moltoni's implication that there may have been some Umbrian influence involved. Perhaps we have to do here not with a general treatment of the cluster *kl*, for example, but with the individualized interpretation of a number of different Latin words. Those for which there was an obvious Oscan equivalent—as in the case of many names—might simply be replaced by the Oscan word without the *c;* other words would retain the proper Latin form. But such an explanation, although it could account for some of the examples, certainly would not account for all.

Instances of morphological interference also occur. Some of them consist of the straightforward application of Oscan endings to Latin words, as in nominative plural *-as* for *-ae*, and third conjugation *-es*, *-el* for *-is*, *-il*. Others, where the form on the inscription is neither Oscan nor Latin, are less clear: *o*-stem nominative plural *-eis* or *-es*, which is neither Latin *-i* nor Oscan *-os*.

In general, it becomes clear that if we are to extract as much information as possible from the treatment of Latin by imperfect bilinguals it is first necessary to lay aside the more remote historical considerations. What happened in prehistoric Oscan is certainly an important background to Oscan as we know it, but it is necessary to avoid the confusion of earlier developments with the state of the language at the time of these inscriptions. And it is equally necessary to delay judgment on the effect of Oscan on Vulgar Latin

and the later Italian dialects until it is clear what Oscan was like at the time it would have been exerting such an influence. For example, the author, with many others, assumes that syncope in modern forms is traceable to Oscan syncope. But for how long did syncope play a role in Oscan? That it began as early as the Oscan-Umbrian period seems evident from the close agreement in detail with the same phenomenon in Umbrian. But the rich development of anaptyxis in Oscan, particularly when it occurs at just the place in the word where we would expect syncope, strongly suggests that the syncopating period was long gone and that the apparent instances of syncopation are but overcorrection arising out of the prevailing habit of pronouncing liquids with an anaptyctic vowel. As noted above, every such instance is in the presence of a liquid. If further investigation, which is needed, proves this surmise to be true, we shall have grounds for wondering whether other claims of Oscan influence on Vulgar Latin may not be based on similar misinterpretations of the Oscan evidence.

WILLIAM DIVER.

ULLA WESTERBERGH, *Chronicon Salernitanum. A Critical Edition with Studies on Literary and Historical Sources and on Language.* Studia Latina Stockholmiensa III. Lund: Almquist and Wiksell, 1956.

This volume provides us above all with an excellent critical edition of the text of a chronicle which was meant as a continuation of Paulus Diaconus' *Historia Langobardorum*, and was composed sometime around 975. The manuscript on which the edition is primarily based (Vaticanus 5001) postdates the time of original composition be some 300 years.

The general principles followed by Westerbergh in establishing the critical edition are praiseworthy. She studies and systematizes lapses from classical Latin and other features, and uses the system as a guide in deciding between variants. In addition she also consults original 9th century documents (especially the *Codex Cavensis*) in order to arrive at 9th century norms.

Pages 222-295 deal specifically with the language of the *Chronicon* and compare it with that of similar Late Latin texts. Without ever stating it in just that fashion, the author shows that the language of the *Chronicon* appears characterized by three features: (1) influence of the contemporary popular speech—such as has

been demonstrated by other investigators of Late Latin; (2) deliberate avoidance of popular features such as lack of the *quod* or *quia* construction for indirect discourse in the narrative, while the same construction is used abundantly in quoting from direct speech (p. 281); (3) occurrence of constructions which are neither Romance nor Classical, but which, since they are used systematically, can not be classed as mistakes, but as features of a linguistic system, such as, for instance, the use of *fore* in "if" clauses (p. 283).

What seems indicated, then, is a general reconsideration of Late Latin after perhaps the 7th or 8th centuries. Evidently it is neither the spoken language of the people as some investigators have over-emphasized, nor is it a Classical Latin with haphazard mistakes. It is a linguistic system in its own right—a vehicle of expression for the educated stratum of the population.

I have always found that the Romance features of Late Latin documents from the 7th century on correspond rather closely to regional linguistic conditions. Thus I feel that Westerbergh's statement that "voiced and unvoiced plosives can stand for each other medially" (p. 225) is not correct. Such a confusion is characteristic of the Late Latin of regions where Latin unvoiced stops became voiced and where, after the voicing (*t*>*d*) and before the simplification of geminates (*tt*>*t*), an unvoiced:voiced contrast (*t*/*d*) did not exist in medial position (cf. *Word* 11, 416-420). Actually Westerbergh gives only examples of the unvoicing of stops, such as *placita*<*placida*, a development for which parallels can be found in the Neapolitan and other Southern Italian dialects (*gravidus*>*gravelo*, etc.). She does not give any examples of the voicing of unvoiced stops, and I rather doubt that there are many in the text. At least I did not find any in a 25-page sample which I scrutinized rather carefully. The original documents of the *Codex Cavensis* which were written at the same time and in the same locality as the *Chronicon* show no examples of the unvoicing of *t* and *p*, and only a few isolated instances of *c*>*g* (cf. de Bartholomaeis, *Archivio glottologico italiano* 15, 262-264, 1901).

Some of the linguistic findings are highly interesting. Westerbergh contends that the replacement of *um* by *o* is more frequent in neuters than in masculine words (p. 237), a fact which she connects with the *-u* ending of masculine words found in some Central Italian dialects, and the existence of the *lu-* article in the Southern Italian area (*ũm*>*o* but *ũs*>*u*?). She also adduces evidence for the Norberg hypothesis that the Romance loss of a

distinct nominative case had primarily syntactical origins and started first in sentences in which the verb is either intransitive or passive. Apparently 75 % of all the instances of loss of distinct nominative occur in such sentences (p. 235). But since we do not really know which type of sentence is most frequent, it would have been more meaningful to compute the percentages of the loss of the nominative within each category of sentence. At any rate, the author's findings demand further research in Late Latin documents, especially the less literary *Codex Cavensis* in which correspondences with Romance developments should be more easily detectable than in the *Chronicon*.

In conclusion, I wish to emphasize that this work is a most valuable addition to our knowledge of Late Latin, and the author's devotion to scholarship and accuracy are to be highly commended.

University of Michigan. ROBERT L. POLITZER.

MORITZ REGULA, *Historische Grammatik des Französischen, II, Formenlehre*. 190 pp. Heidelberg: Carl Winter, 1956.

The corresponding *Lautlehre* was reviewed in Volume 11 of this journal (pp. 466-468), and most of what was said there could be repeated here: this is a handy and well-informed presentation of historical French morphology; the little volume offers far more than its format and bulk would seem to promise. On the debit side, there are more misprints and material errors than is condonable in a handbook for beginners. Here again, as in the *Lautlehre*, time-depth is largely ignored. But since the nature and causes of changes are better known in morphology than in phonology, it will as a rule be easier here to reestablish time perspective with the help of the sketchy data we are offered. Yet a sprinkling of dates would have been quite welcome, as when the following analogical masculines are presented promiscuously: *fé* from *fée*, *machin* from *machine*, *puceau* from *pucelle*, *m'ami* from *m'amie*, *emperier* from *emperiere*.

ANDRÉ MARTINET.

HELMUT GNEUSS, *Lehnbildungen und Lehnbedeutungen im Altenglischen*. viii+184 pp. Berlin: Erich Schmidt Verlag, 1955.

Although a great deal has been written on loan words from Latin in Old English, much less attention has been given to the

more subtle types of linguistic borrowing which are dealt with in Helmut Gneuss's dissertation. The only previous attempt at a classification by types of Old English borrowings from the Christian-Latin vocabulary was made by H. S. MacGillivray in his Göttingen dissertation of 1902;[1] since his book was a study in cultural history rather than in linguistics, he did not pursue the subject into details. In 1926 Samuel Kroesch published a study of "semantic analogy" and its effect upon borrowings,[2] and three years later, in an article on "Semantic Borrowing in Old English,"[3] he applied his principles to the specific problem of Latin loan material in Old English. A number of other earlier studies are discussed by Gneuss in his survey of previous scholarship (3-15), but in the Old English field MacGillivray and Kroesch seem to be the most significant forerunners of the present volume.

The terminology and method of Gneuss's study are adapted from two studies of the Latin influence on Old High German by Werner Betz.[4] Following Betz, Gneuss classifies the loan material ("Lehngut"), other than loan words proper, under the following four heads :

Loan-formations ("Lehnbildungen"), which can be further classified into *loan-translations* ("Lehnübersetzungen"), *loan-adaptations* ("Lehnübertragungen"), and *loan-coinages* ("Lehnschöpfungen");
Loan-meanings ("Lehnbedeutungen");
Loan-expressions ("Lehnwendungen");
Loan-syntax ("Lehnsyntax").

The last two of these categories, loan-expressions (such as OE *on woruld worulde* for the Latin *in seculum seculi*) and loan-syntax (as found, for example, in the Old English dative absolute construction, *him andweardum* modeled on the Latin *eis presentibus*) are relatively infrequent in Old English and receive only brief treatment. The major part of Gneuss's book, then, deals with loan-meanings and with the three types of loan-formations.

[1] *The Influence of Christianity on the Vocabulary of Old English*, Part I (Halle, 1902); see especially pp. 148ff.

[2] *Language*, 2: 34-45 (1926).

[3] *Studies in English Philology, A Miscellany in Honor of Frederick Klaeber* (Minneapolis, 1929), pp. 50-72.

[4] *Der Einfluss des Lateinischen auf den ahd. Sprachschatz: I. Der Abrogans* (Heidelberg, 1936); *Deutsch und Lateinisch: Die Lehnbildungen der ahd. Benediktinerregel* (Bonn, 1949).

Without going into detail, I will summarize these categories in the next two paragraphs.

Loan-meanings (the subject of Kroesch's earlier investigation) are a result of the transference, by analogy or substitution, of a Latin meaning to an already existing Old English word. Examples of analogical transference are OE *lunge* "language" <Latin *lingua*, OE *ðeode* "heathen" <Latin *gentes*, and OE *fæstnes* "firmament" <Latin *firmamentum;* in each of these cases, the Latin and the Old English word had similar literal meanings, but the Latin word also had an extended, figurative meaning which was transferred to the Old English word. Examples of substitution are OE *cniht*, for Latin *discipulus*, and OE *halig*, for Latin *sanctus*.

Unlike loan-meanings, which always involve previously existing words in the borrowing language, loan-formations result in new words, either compounded or derived. Examples of loan-translations, by far the most frequent of the three types, are OE *anboren*, for Latin *unigenitus*, and OE *ðrines*, for Latin *trinitas*. Examples of loan-adaptations (which differ from loan-translations in not showing a one-to-one correspondence with the elements of the Latin word) are OE *restedæg*, for Latin *sabbatum*, and OE *forðrestnis*, for Latin *contritio*. Finally, in loan-coinages, the least frequent of the three types, we have paraphrases of the Latin words, rather than complete or partial translations; examples are OE *fagwyrm*, for Latin *basiliscus*, and OE *cwildeflod*, for Latin *diluvium*.

Nearly two-thirds of the book is devoted to an application of these categories to the loan materials in the Old English interlinear psalter glosses, and particularly to those in the oldest of the glosses, the Vespasian Psalter. In all, 693 Old English words are studied, and are conveniently listed in one of the indexes; of these, 533 are loan-formations and 160 are loan-meanings. The total for the Vespasian Psalter itself is 292 words, including 201 loan-formations and 91 loan-meanings. Probably not all users of the book will agree with all of Gneuss's dispositions of individual words; in fact, he specifically warns us of the frequent difficulties of analysis, particularly in the distinction between loan-translations and loan-adaptations. But there can be no question that in this book, for the first time, we have a thorough and systematic classification of the various types of Latin loan materials in Old English, which can be used in the analysis of loan materials in other Old English texts, in place of the rather impressionistic interpretations which have been the rule up to now. Hearty

congratulations are in order, not only to the author but also to his teachers in the Free University of Berlin.

Columbia University. ELLIOTT V. K. DOBBIE.

DANIEL JONES, *An Outline of English Phonetics.* Eighth edition, completely revised and re-set. xx+378 pp. Cambridge: W. Heffer & Sons Ltd., 1956.

DANIEL JONES, *Phonetic Readings in English.* New and improved Edition (re-set). xx+90 pp. Heidelberg: Carl Winter, 1956.

Daniel Jones, a giant figure in the world of phonetics, has been very busy during the last few years revising, enlarging, improving and in general tightening up the earlier editions of his many works.

His *Outline of English Phonetics* was first published in 1918, was re-written as a third edition in 1932 with the fourth to seventh editions being mainly reprints of the third. The present book is the eighth edition.

Phonetic Readings in English, intended for use by foreign learners of English in conjunction with the *Outline*, was first published in 1912 and went thru many subsequent printings. It, too, has now been brought up to date.

As the publisher's blurb on the first book jacket states: "The book treats of English (Southern British) pronunciation primarily from the point of view of the foreign learner of English. The various features of English pronunciation are described in minute detail, the chief mispronunciations heard from foreign people are explained, and methods are given for enabling students to pronounce correctly..."

Workers in the language-speech field are likely to approach these books from several different angles, depending on where they happen to be standing at the moment of impression.

Persons trained in IPA notation and in the application of it to the teaching of speech will not be critical of the books.

Structuralists of the various persuasions, however, will quite naturally find themselves re-casting everything in the newly developed molds that have become second nature to them and will come up with many unanswered questions, doubts, and dissatisfactions.

Indeed, there are many who may feel an acute sense of frustration in seeing the "old master" continue to steer his own course,

seemingly little affected or perturbed by the structuralist currents eddying about him.

And yet, only if Jones' unswerving devotion to the problems of phonetic writing and to the teaching of pronunciation is accepted and understood as the legitimate right of the individual to choose his own theater of operations can one even begin to judge his work.

For, while it is not difficult to point out flaws in the scientific side of Jones' thinking, it is really on Jones' record as a teacher that his work must stand or fall.

But there is no danger whatever of any but an affirmative verdict. The *Outline* speaks to the teacher in a language that he understands. For he knows that, in the art of teaching, empiricism is often more significant than pure science.

It is to books like the *Outline*, then, that he will look for such basic principles as those dealing with ear-training, cardinal vowels, assimilation and similitude. Jones makes clear that the learner is not expected to learn vowels by "simple imitation," nor even mainly from a knowledge of their organic formations, but rather by learning to estimate by ear the extent and nature of acoustic distances from one vowel to another.

On the other hand, the learning of consonants is facilitated by the knowledge of the facts contained in an organic description and by the application of them to one's own speech organs. Tactile and muscular sensations rather than sensory control from the ear play the biggest role in this case.

It is in connection with the learning of vowels that Jones' work in phonetic trancription comes to the fore. Since one of the main goals in using transcription to teach pronunciation is the establishing of visual associations between the symbols and the desired sounds, Jones' discussion of his experience with the various types of symbols in teaching students of varied backgrounds cannot help being one of the most suggestive parts of the book to any classroom teacher of speech or of foreign languages.

But Jones does not neglect the reality of English orthography. Aware that the literate learner of the language must of necessity come to terms with the spelling system, he relates each phoneme with the letter (or letters) that usually represents it at the same time that he takes it up in phonetic detail.

Altho the *Outline* is certainly not a "do-it-yourself" book nor a simple classroom manual, it is a book that should be the vade mecum of every teacher of speech and of every serious student of English pronunciation.

A need that has yet to be filled in this country is an American counterpart to the *Outline* as well as practical manuals and phonetic readers embodying its basically sound pedagogic approach. The growing importance of the American variety of English would certainly seem to call for an accelerated effort on the part of American phoneticians to catch up with the Jonesian output.

Nothing in this review should in any way be taken to deny the value of Daniel Jones' work for linguistic analysis per se. Quite the contrary; all of what Jones' has to say must be reckoned with by any serious student of the English language.

This reviewer for one believes that Anglicists who neglect to study closely Jones' detailed discussions of stress, vowel tamber and prominence in the *Outline* and elsewhere on factional or ideological grounds do so at their own peril.

It has been noted elsewhere that Jones has not kept up with the burgeoning structuralist literature of recent years, but let us not delude ourselves that his failure to incorporate more than a token number of structuralist observations in his own scheme of thinking is due to mere benightedness.

From his comments on the structuralist works that he has read, it is clear that the "old master" is simply not only unaffected but unconvinced.

Without losing sight of the limitations that Jones seems to have set for himself, let us at least be objective enough to admit the possibility that we have not done enough to convince him.

The City College of New York. MARSHALL D. BERGER.

DANIEL JONES, *Everyman's English Pronouncing Dictionary.* xlv+538 pp. London: J. M. Dent, New York: E. P. Dutton, 1956.

What we are offered, under this new title, is the 11th edition of a classic, Daniel Jones's *An English Pronouncing Dictionary* completely revised, enlarged, brought up to date, and reset. Close to 2000 new words and names have been included; new pronunciations of old entries have been listed; some variant pronunciations have been promoted from second to first rank, others demoted from first to second; more extensive use has been made of the hyphen as a mark of syllable separation; a welcome distinction has been established between the transcriptions of the second half of *windier* and *reindeer*, *rescuer* and *rest-cure;* foreign

variants have, for the first time, been given for languages other than French and German; large sections of the Introduction and Explanations have been rewritten, and a glossary of phonetic terms has been added at the end of the book. Let's hope that the new frame in which this excellent book is now published will make it available to still wider sections of the international public.

ANDRÉ MARTINET.

MAX NIEDERMANN, *Balto-Slavica* (Université de Neuchâtel, Recueil de travaux publiés par la Faculté des Lettres, 27th fascicle). 215 pp. Geneva: Droz, 1956.

This volume is a collection of the most important contributions of Max Niedermann to the field of Baltic and Slavic philology. According to the editor, G. Redard, the texts are reproduced in their original form, with the exception of a few of the author's comments made in the margin of his own offprints and the normalization of the orthography in accord with modern practice. Altogether there are eleven articles devoted to (1) the names for the stork in Lithuanian (pp. 9-32), (2) names for the potato in Lithuanian (pp. 33-116), (3) various syntactic similarities which Lithuanian shares with other Indo-European languages (pp. 117-122), (4) forms of address in the standard language (pp. 123-135), (5) Lithuanian terms for a 'dapple-gray horse' (pp. 136-137), (6) etymological studies dealing with names of tools and trees (pp. 138-153), (7) O. C. S. *bridŭkŭ* (pp. 154-156), (8) Lett. *zandarts* 'stork' (pp. 157-158), (9) The origins of Lithuanian linguistics (pp. 159-173), (10) back-formation in Lithuanian (pp. 174-178), (11) studies in Lithuanian word-geography (pp. 179-196).

The chief value of the volume lies in the exhaustive treatment of the lexical items. For example, the study of the names for the potato in Lithuanian is, to the best of my knowledge, the most detailed treatment in Baltic linguistic literature, although Senn's study *Lithuanian Dialectology* (Supplement I of the *American Slavic and East European Review*) does contain several additional words. Since this book is equipped with an index (pp. 197-214) it can be used as a supplement to an etymological dictionary.

One is inclined to doubt the value of comparing similar syntactic features common to various geographically widely separated Indo-European languages. For example, we find (p. 117) that in Old Persian, Latin, Pali, Lithuanian, Polish, Czech, Serbo-

Croatian and Russian the expression 'to die a natural death' is rendered by an expression which literally translated would be 'to die one's own death.' To the reviewer this agreement seems to show only that similar constructions in the Eastern European languages might be related, but here probably borrowed from one to another rather than from common I. E. inheritance. On p. 117 Niedermann has noted that in Latin and Lithuanian sometimes the substantives *husband* and *wife* are omitted and the possessive pronoun is used to denote 'my (your, his) wife or husband.' It does not appear likely that this is I. E. inheritance.

In the article 'Gli inizi della linguistica lituana' we find (p. 163) that, "La grammatica di Schleicher è rigorosamente descrittiva, perchè l'autore, di proposito deliberato, s'è astenuto da qualsiasi comparazione con le altre lingue indoeuropee..." American trained linguists might quarrel with the above statement, but we must remember that it was written in 1931.

In any case, whatever Niedermann's evaluation may have been, this article does contain much basic information about the early history of the study of Lithuanian.

To conclude, this book is valuable as a repository of facts, particularly lexical, but the interpretations should be carefully examined before one accepts them.

University of Kentucky. WILLIAM R. SCHMALSTIEG.

GUNNAR HERNE, *Die slawischen Farbenbenennungen; eine semasiologisch-etymologische Untersuchung.* (Publications de l'Institut Slave, IX). 147 pp. Uppsala, 1954.

The author has confined his treatment in this book to those color names which are common to several Slavic languages. Words of more limited currency have been deliberately disregarded unless they have become the most general designation for a color, e.g. Polish *niebieski* 'blue'. Words like Russian *malinovyj* 'raspberry (-colored)' whose derivation from a substantive is obvious, are also left out. The study was chiefly aimed at delimiting which types of color names are expanding and which are decreasing. The problem, as Herne sees it, has two aspects: adjectives denoting colors can be original, derive from substantives, or else go back to words meaning 'bright'; on the other hand, certain color adjectives are applied to certain objects only, as hair, eyes, etc., whereas others are used freely. In both cases the author's aim

is to determine whether one of the types mentioned may be considered primordial.

In his treatment of the first problem, the author focuses his attention on the adjectival derivation from substantives, while the relation between names of colors and words denoting 'light' and/or 'darkness' are somewhat neglected. Only in passing do we find references to the latter, as when it is said that *sivŭ* and *sinĭ* may be derived from a word with the meaning of 'dark' (120): this, by the way, is a slip since **sivŭ*, *sinĭ* together with *sizŭ* should rather be traced back to a stem denoting 'light', found in Russian *sijat'* 'to shine' and elsewhere, as is correctly indicated on p. 83. The evolution of Cz. *blankytný* 'blue' from MLat. *blanketus* 'panni genus albi', Pol. dial. *jasny*, Cz. dial. *světlý*, both 'blue', provides us with convincing parallels. The use of *bělŭ* 'white' in the meaning of 'bright, light' in R. *belyj den'*, *belyj svet*, and the like (106) may be paralleled by the passage from a Czech dictionary of the 14th century, the so called *Vokabulář gramatický: bělosvět-gallaxia* (V. Flajšhans, *Klarel a jeho družina*, sv. 1, Prague 1926, p. 17). More examples of this kind are available, and certain patterns of development might be established. The assumption of Havers (Neuere Literatur zum Sprachtabu, *Sitzungsberichte* of the Vienna Academy, 223, 5, p. 118), going back to Vendryès, that IE color adjectives differed only in their initial consonants (**pel-*, **bel-*, **bhel-*, **kel-*, **gel-*, **ghel-*, **g^{w}hel-*) remained in all probability unknown to Herne. He does not investigate whether or not this assumption has any bearing on (Common) Slavic. Old Rus. *izekrē* 'blue eyed' (quoted on p. 61, but not analyzed since it is a monolingual word) may represent the same problem if the word actually goes back to the CS1 stem *zĭr/zor-* 'to see' as Sreznevskij suggested (*Materialy dlja slovarja drevnerusskogo jazyka*, I, 1085). It affords another interesting example of specialization in semantic development which results in a word denoting a color.

Regarding the derivation of color adjectives from substantives, particularly those denoting animals, Herne would appear to favor the thesis that the adjectives usually are secondary. However this contention seems to be premature if applied to Rus. *goluboj* 'light blue'. The case should not be identified with that of *voron—voronoj*, *zoloto—zolotoj*, etc. The semantic connection, which is obvious in the latter examples, is not perfectly clear in the case of *goluboj* as linked to *golub'* 'pigeon'. Besides, there is also a structural difference: whereas, in other cases, both adjective and

substantive belong to the same stems, we have different stems in *golub'* and *goluboj*. (The point of view contrary to that of Herne is now expounded by Moszyński, *Język polski*, XXXV, 189ff.). As to the stem **rūdh-so*, the possibility of connecting it with the name of the lynx, *rysĭ*, should be investigated.

Herne does not reach definite conclusions as to whether color adjectives normally tend to restrict the sphere of their applications from general to object-bound, or whether the reverse is true. According to him Cz. *hnědý*, Ukr. *hnidyj* 'brown' is the only case of generalization of an originally object-bound adjective (120). As to Ukrainian, Herne was here misled by his native informant who presented him as general Ukrainian the meaning which the word has in W. Ukr. dialects influenced by Czech. A glance in Hrinčenko's dictionary would have shown that *hnidyj*, in its application to objects other than horses, is restricted to the westernmost Ukrainian dialect of *Lemky*. From only one Czech word of dubious etymology ("Vielleicht ein Lehnwort", says Herne, but from what language? p. 61) generalization is certainly hasty.

An analysis of IE and Slavic taboos would substantially contribute to the question of what was primary in the meaning of color adjectives—their "object-boundness" *(Gegenstandsverbundenheit)* or general usage. Color adjectives, as is well known, are sometimes used to replace animal names like 'the red one', 'the gray one', 'the brown one' instead of 'fox', 'wolf', 'bear', etc. This can entail the narrowing of the sphere of application for the adjectives concerned. On the other hand, the tabooing of color adjectives themselves is not implausible. This is Havers' opinion (*Neuere Literatur...*, pp. 48, 119), which is probably correct. Yet other views were expressed too, denying the role of taboo (Cf. L. Weisgerber, Das Problem der inneren Sprachform und ihre Bedeutung für die deutsche Sprache, *Germanisch-romanische Monatsschrift*, XIV, 252), and it would be gratifying to investigate the problem and the controversy with Slavic material. Is it, for instance, an accident that *sinĭ* was crowded out of Polish, except in the "sombre" or "gloom" connotations, as applied to corpses, sores, bruises, etc., while Eastern Slavic, which possesses at least for a part of these connotations the adj. *bagrovyj/bahrovyj*, has fully preserved *sinĭ*? These interesting questions which might find their answers if the SYSTEM of color adjectives were considered are not even hinted at in Herne's work.

The author's general conclusions are cautious. He expresses the supposition that the modern system of color denotation in Slavic may be traced back to a more reduced one consisting of three colors (red, yellow, green), and the latter in its turn to that of red and yellow-green only. But he adds immediately that these conclusions are very hypothetical, and he underlines the word *very* (121). Otherwise, a conclusion seems to be implied that semantic development of color adjectives is general and always the same, so that the different and mutually opposed processes of generalization in meaning and of its restriction are going on simultaneously and uninterruptedly. There are no grounds to deny in principle such a conclusion if it is actually implied. Still in Herne's study it is rather imposed by the method applied than resulting from the evidence. Says Herne: "Die gleiche oder ähnliche Bedeutung eines Wortes in mehreren Sprachen oder Sprachfamilien berechtigt uns, dieselbe Bedeutung auch für die Ursprache anzusetzen" (21f.). When different languages employ the same word in various meanings Herne follows the statistical majorities. If the number of languages are equal he is in a quandary, cf. his difficulties in the explanation of *gnĕd-* as to its original meaning, because he supposes that the word is used in two languages in general meaning and in two others in specialized (60). His comparative and even statistically comparative method limit the scholar to the quantitative changes and make him unable to reveal qualitative changes in the Slavic color names. It is typical of the positivism of late 19th-century linguistics.

Another belated repercussion of this approach seems to appear in Herne's idea of constant perfection in language evolution and of modern European languages as the best developed, e.g., when he says that "in Bezug auf die Farbenbenennungen kann eine Sprache als völlig entwickelt gelten, wenn sie für sechs Hauptfarben allgemeine Bezeichnungen hat, wie dies in den meisten heutigen europäischen Sprachen der Fall ist" (19). Compare with this for example L. Weisgerber's idea that the modern European color system in its unity could have arisen as a result of borrowings (I would prefer to speak of belonging to the same *Sprachbund!*) and not of progressive perfection inherent in the linguistic development as such: and that the color system of ancient Greeks and Romans, though very different, was not more backward or imperfect (*Muttersprache und Geistesbildung*, Göttingen, 1929, 81f.).

His semantic analysis lacking the notion of qualitative changes and the notion of system, Herne's work proves to be a collection of word histories and etymologies (if one wants to distinguish the two). From this point of view, however, the book is, despite some shortcomings, an accurate and useful work. In studying word histories and etymologies Herne takes his point of departure in the modern Slavic languages and not in the data of old texts. The latter are not discarded, but they are discussed in the light shed by the modern and more easily controllable material. Words are considered in the spheres of their use (linguistic genres), not in abstraction. This enables the reader to see how widespread the words in question are and what are the limitations in their use in a given language. Derivational peculiarities and connections are thoroughly scrutinized. This makes it possible to see how deeply the given word is rooted. Herne endeavors to follow the connection between the word's history and history of objects *(Wörter und Sachen)* and that of society. Cf. his digressions on the production of red dye or the elucidation of what is the part played by blue color in Slavic rural life (37 ff., 87, 97). Incidentally, the latter accounts for the lack of a native word for 'violet' in Slavic.

Herne takes into consideration the vocabulary of all the Slavic languages, except, unfortunately, Macedonian. Slovak and Sorbian are treated insufficiently, but some languages usually explored superficially or inadequately are used by Herne broadly and in most part accurately, in particular Ukrainian and Belorussian. (But was it necessary to preserve in Belorussian the spellings of various and in this respect extremely variegated sources?). Herne has built his work chiefly on available dictionaries of the Slavic languages. He was able also to utilize the data of still unpublished parts of the *Rječnik* of the Yugoslav Academy and of the *Slovník staročeský* of Gebauer. To construct a work on dictionaries is fraught with dangers. One of them is the risk of operating with ghost words. Herne has shunned this danger by checking up his material with his native informants. This also enabled him to establish whether a given word is used broadly or is limited to certain linguistic genres. But the other consequence of adhering to dictionaries, namely aligning of linguistic facts which actually do not belong to one line of development—weighs heavily in some parts of Herne's book. E.g., the material of the Serbo-Croatian *Rječnik* presents data from Old Serbian and from the language of the 16th-17th century

cities on the Adriatic littoral along with modern Serbo-Croatian. These data did not grow up on the same dialectal base and there often is no continuity among them unless in the framework of the artificial literary language. The main sources for Middle Polish and the Polish dialects, the dictionaries by Linde and Karłowicz, are often unreliable, as recent investigations have shown (E.g. K. Nitsch, *Studia z historii polskiego słownictwa*, Cracow, 1948: W. Kuraszkiewicz, Rad bywa smárd, gdzie rájtarká, *Język polski*, XXXIV, 1954). But it is particularly difficult to use correctly Sreznevskij's Old Russian dictionary. First, it is incomplete because it does not encompass the publications of old texts during the last half century. Still more important is the fact that Sreznevskij's collection of data, although entitled an Old Russian dictionary, actually aggregates the words belonging at least to three linguistic units: Old Church Slavonic in its Eastern Slavic recension which served as a common literary language for all Eastern Slavs: dialects underlying later Russian; and dialects underlying later Ukrainian. To ascribe all this to Russian and to construct a straight line of development from Sreznevskij's data to contemporary Russian means to build an imaginary edifice. So **rūdrŭ*, characterized (24), as Old Russian and modern dialectal (Northern) Russian, is evidently only Old Church Slavonic and Northern Russian. The oldest examples of **xěd-* in Eastern Slavic are not Russian but Old Church Slavonic (113). The statement that Rus. *krasnyj* acquired its modern meaning of 'red' in the 15th century is arbitrary. It is based on the lack of examples for this meaning in Sreznevskij prior to that time. But, in fact, no examples of this adjective in proper Russian are cited at all, and one cannot draw conclusions *ex silentio!*

The oldest dictionaries are unfortunately not utilized. They could furnish valuable data unavailable otherwise. A few examples will suffice. The Czech dictionary of the 14th century, by Claret, brings this enumeration of color names: "Viriditas *zelenosl*, sed griseitas dato *sserosl*. Et rubedo *czervenosl* fuscitudo quoque *smyedosl*. *Bledosl* palitudo, *narissenstvie* sit flavitudo. Flaveitas *modrosl* dic, citrin<itas> quoque *rusosl*. Glaucitas est *zlulosl* carnalitasque *tyelesnosl*" (Flajšhans, 160). Ukrainian dictionaries of the 16th and 17th centuries disclose what the contemporaries judged as Ukrainian and what as Church Slavonic. E.g., *Lexis* by L. Zyzanij (1596) qualifies *bahrov<yj>* as Ukrainian, and this would preclude Herne from putting dial. *bahryvyj*, *bahryj* in the first place in his paragraph on this word in Ukrainian (48).

Also *sidyj* in Ukrainian is either archaic or dialectal, and was so as early as the 17th century: at that time *Synonima slavenorosskaja* attributes *sědina* to Church Slavonic and quotes as its Ukrainian correspondence *syvyzna* (P. Žytec'kyj, *Narys literaturnoji istoriji ukrajins'koji movy v XVII vici*, Lvov 1941, 184: cf. Hrinčenko, 2,635).

The central part of Herne's study are his etymologies. In his etymologizing he is cautious and often prefers to leave questions open rather than to venture too bold hypothesis. Old attempts are rejected and no new ones suggested for *bagŭrŭ* (See F. Sławski in *Rocznik slawistyczny*, XVI, 1, 85), *sizŭ* (50, 84), etc. Nevertheless some etymologies accepted or suggested by the author seem to be debatable, because sometimes he is guided only by exact phonetic correspondences, disregarding the contradictory social and cultural background. It is hardly plausible to deduce South Slavic *sur* 'ash-gray' from Rumanian only because no form with *u* is attested in modern Turkish. The use of the word in Serbian of the 17th century and particularly in Serbian folklore indicates ancient borrowing which hardly could have been from Rumanian: on the other hand, we know too little about Old Turkish or Proto-Bulgarian vocalism to reject the old etymology of Miklosich and Korš only for a discrepancy in the vowel *u*∾*o* (117).

For the same reason Herne denies Turkish mediation for Persian *bor* used in Russian (and East and West Slavic) as *buryj* 'brown' (65), although a direct borrowing from Persian into Russian, not to speak of Western Slavic, is completely implausible.

The most venturesome is Herne's attempt to consider **čĭrm-* 'red' as an old Iranian loanword to be added, then, to generally acknowledged Iranisms in Slavic as *toporŭ* 'axe' or *sobaka* 'dog'. Ample digressions are devoted to this etymology. Some of them are unnecessary, as e.g. that on the chronology of first palatalization in Slavic and its later repercussions. This may be eliminated by the assumption that the Novgorod area had been colonized by the Slavs by the 6th century A. D. If, as Herne assumes, the borrowing occured in the 3rd to 1st century B.C. (42f.), a mediation of Iranian "Reisende und Kaufleute" is superfluous because at that time the Sarmatians controlled the steppes North of the Black Sea. The main objection against Herne's view results from an insight into the distribution of the word and its connotations in the Slavic languages. A color meaning for this adjective was typical only of Old Church Slavonic, from which it penetrated literary

Eastern Slavic. Hence Sreznevskij's examples. Western Slavic, Slovenian included, knows the word only in the names of animals and with the meaning of 'yolk'. Yet, in the case of a borrowing from the Iranians, who alledgedly brought to the Slavs fabrics painted in this color and/or technics of their painting, it would be natural to expect that the prevailing connotations in Slavic would be those with clothes and color. For this reason Herne's attempt at augmenting the number of old Iranian loanwords in Slavic is a failure like so many such attempts made previously. It is more reasonable to adhere in the case of **čĭrm-* to the old theory of an interplay in the stem formants *-v-*∾*-m-* as advanced by Miklosich and others (34f.). See also F. Sławski, *Słownik etymologiczny języka polskiego*, p. 119.

Other etymologies by Herne are not so original. But a lesser or greater number of acceptable etymologies does not justify the appearance of a book on color names as long as they are not shown in their interrelations as parts of a system. We would like to know whether color names in Slavic make up a coherent group in which a change of one of the members results in a change of the whole system. Is for example Bréal's principle of *répartition* (*Essai de sémantique*$_2$, 28ff.) valid for the group of lexemes denoting colors? The opposition of 'black' and 'white' is evident, but what about other color names? Are there any intrinsic regularities within this group of words?

To discuss this problem one should operate within the framework of one language, with all adjectives denoting colors under scrutiny. Herne deals with few words selected because they are common for all or several Slavic languages. With this approach only some regularities of external nature may be disclosed. Such is the history of **reudh-* in Slavic. Herne demonstrates how it was gradually limited in most Slavic languages to restricted object-bound use as applied to hair of men or animals. The reason was mainly the diffusion of the words with the stem **čĭrm-*∾**čĭrv-* due to new techniques of painting fabrics. Another factor hinted at by Herne but not examined more deeply could be the homonymy with *ruda* 'ore'. The possible influence of another semantic ramification of the stem, to wit *ruda* 'blood' (see e.g., Dal', Hrinčenko *s. v.*) should also be investigated. But in all these cases we deal with changes which are conditioned from outside.

Having found no internal regularities in the development of color names and, presumably, still endeavoring to reveal some, Herne tries to appeal to what is called natural colors. But this

is a false issue. The logic of reality here as everywhere is not necessarily the logic of a language. As Bréal stated it: "Il n'est pas douteux que le langage désigne les choses d'une façon incomplète et inexacte" (*Essai*, 191). The distribution and the limitations of colors are determined by a language and not in a language by reality. Havers indicates that taboos for fox in some languages are 'red', in others 'yellow', for toad—'red' and 'brown' (*Neuere Literatur...*, 48f.). Slavic designations of various hues of blond hair belong usually to the domain of 'yellow' or even 'white' and not to that of 'red' as they do in Germanic. For this reason the place of *rusŭ* 'blond' (28f.) is wrong in Herne's work, and his remark that *al* 'scarlet' cannot denote color of hair (50) is superfluous (but see Sloven. *rdèč* and the correct general formulation on p. 56). Ultimately, the collapse of the approach from the stand-point of natural colors is demonstrated by the fact that Herne devotes special chapters to black, white, gray, and brown which do not belong to the spectrum, while no attention is paid to the violet color, which does.

Finally some minor shortcomings or lapses may be pointed out.

The author rightly excludes verbs from his survey, but he is wrong in denying the use of sentences like Rus. *parus beleet* in vernacular and omitting the sentences of the type *parus beleetsja* (20). Old Church Slav. *prisno* is not "früher" but 'always for ever' (27). In the survey of etymologies of *ruměnŭ* (28) the opinions of A. Vaillant on the suffixes *-ěn-*∾*-ęn-* in his *Grammaire comparée des langues slaves*, I, 147, and L. Sadnik (*Mélanges F. Ramovš*, *Slavistična Revija*, III, 3-4) should be mentioned. Ukr. *ryžyj* is characterized as being applied to the hair of man, this with a reference to Hrinčenko. But Hrinčenko's only example refers to a calf. Ukr. *červonyj*, Blr. *čyrvony* 'red' are not borrowed from Polish (46). They and Pol. *czerwony* are rather a result of complex and combined interrelations of the three languages; see my article in *Word*, 8.340f. (1952). In addition, in the Blr. word *y* cannot be due to its unstressed position. Herne's attempt to deduce SCr. *plâv* 'blue' from German *blau* is unpersuasive. German influence in Serbo-Croatian of the 14th century is insignificant, except in a few terms of mining industry, of which *plav* is not one. All the examples in the *Rječnik* of the Yugoslav Academy (X, 24) show no connection with mining terminology, but point to a possible gradual shift of connotation from 'yellow' through 'yellow-gray' and 'hazy' to 'blue'. Cf. especially: "Marga u boji... jest ili plava ili sivasta ili bjelatasta ili žuta." Pol.

blawatny 'blue' developed in all probability not from *blawat* 'kind of silk', but from *blawat* 'Centaurea cyanis' as Vasmer correctly states in his *Russisches etymologisches Wörterbuch s.v. blavatka*. The fabric in question was of various colors, cf. in Ukrainian of the 17th-18th century *běloho... blavatu, červonoju blavatnoju materijeju* (Je. Tymčenko, *Istoryčnyj slovnyk ukrajins'koho jazyka* I, 1930, p. 94. Incidentally, this important source is not used by Herne). The contention that *blakytnyj* in Ukrainian is used only when speaking of the sky (93) is wrong. A glance at any dictionary would show other uses. *Hranatovyj* 'blue' is alien to Standard Ukrainian and known only dialectally in the Western Ukraine. To South Sl. *mor* 'dark blue violet', Sloven. *mor*∾*mûr* must be added (95). As to *bronŭ*, the material (106) is dubious. It remains unclear whether the word belonged to Old Russian, as well as its connection with Russ. dial. *brunel'*, Ukr. *brynily*, etc. The information of Lubynec'kyj as in many other cases is incorrect when he says that Ukr. *čalyj* 'grey brown' can refer to fabrics.

The author's bibliography as to general treatment of color names in linguistics could be considerably enlarged. E.g., works by Havers, Weisgerber, and Bréal quoted in this review are not cited in the bibliography compiled by the author. It is particularly regrettable that the vast bibliography by Skard (*Proceedings of the American Philosophical Society*, 90, pp. 163-249) was not taken into account. Although primarily concerning the use of colors in literature it contains also a great many items important for a linguist.

Attention has been concentrated in this review mostly on what is debatable, insufficient or erroneous in Herne's work. But as a whole the book has its merits. The bibliography used in it is imposing, the careful approach commands respect, the material collected is in most part reliable, the tendency to find regularities in linguistic development, although so far lacking a proper method, is to be welcomed. The composition of the book is transparent and concise, the scope of the data large. The study does not bring much which is new but it is useful as a confrontation of views presented hitherto and it stimulates further research.

To my knowledge, this is the first book by Herne. It is a promising start. One would wish that it would not also become Herne's last book, as has happened for unknown reasons to many a Swedish Slavicist after the publication of his first book.

Columbia University. GEORGE Y. SHEVELOV.

STANISŁAW BĄK, *Gwary ludowe na Dolnym Śląsku, Cz. I, Głosownia.* (Poznańskie Towarzystwo Przyjaciół Nauk, Prace Komisji Filologicznej, Tom XV, Zeszyt 1). 129 pp. Poznań, 1956.

Studies in dialectology initiated and led by the old master of Polish linguistics, K. Nitsch, thrive in Poland more than in any other Slavic country. Poland can boast a greater number of monographs on separate dialects published after the World War II than Russia, the Ukraine and Czecho-Slovakia, and she is the only Slavic country in which the history of the national language is about to be treated as historical dialectology.

Bąk's book encompasses the speech of 21 villages in the districts of Syców, Oleśnica, Namysłów, Oława, Brzeg and partly Kępno (Germ. Perschau, Öls, Namslau, Ohlau, Brieg, Kempen), East of Wrocław. It joins the other recently published book, by P. Gołąb, *Gwara Schodni i okolicy* (Wrocław, 1955), which describes the dialects of the Opole (Oppeln) district, bordering upon the area covered by Bąk from the Southeast. Bąk's study is conceived as divided into three parts: phonology, morphology, and texts recorded from the speakers. A map, badly needed in the now published part I, devoted to phonology, will be added to one of the following parts.

The territory of Northeastern Silesia is of a special interest linguistically for several reasons. It is Polish without being influenced by Standard Polish, for it has been torn away from Poland since 1163 (up to 1945). From 1327 to 1763, roughly, it was politically and culturally connected with Bohemia. For the significant, albeit often overestimated, Czech influences on Polish it was one of the most important channels. One can hardly imagine any influence of spoken Czech on Polish unless it was conveyed via Silesia. Eventually these dialects proved to be on the utmost outskirts of the Polish linguistic area, practically submerged in the German element and gradually dissolving. This made them invaluable for studies in the problems of bilingualism. On the other hand, because of their destiny, the task of a dialectologist in exploring them now is not only to record the material but to a certain extent to reconstruct the original make-up of the dialect from the later distortions, as was the case of the extinct Great Polish dialect of Chwalim, recently investigated by R. Olesch (*Abhandlungen* of the Academy in Mainz, 1956, 6). In addition, there is no time to lose because after the post-war resettlements the scanty dialect speakers have come into destructive

contact with Polish dialects of numerous new settlers. Under the impact of these influences the last hallmarks of the original Low Silesian dialects are vanishing faster than ever before.

Bąk's book does not deal directly with all these problems. After giving a survey of how the Germanization of the area proceeded, it only presents the material classified separately for reflexes of phonemes of "Old Polish." What exactly is understood here by Old Polish is nowhere explained. The choice of this background is not the most convenient, not even to facilitate references. E.g. the reflexes of *ě* are dispersed in the chapters on *a* and *e:* there is no chapter on reflexes of **tŭrl*, **tŭlt*, **tĭrl*, **tĭlt*. Methodologically it seems to imply that in the author's opinion there was some time when Polish was uniform, without any dialectal divisions, which is implausible.

Bąk is interested in fixing the scanty evidence of the vanishing Low Silesian dialects, and in this solely. Bilinguality of the speakers is only a hindrance for him. In principle he aims at recording only those features which are Old Silesian. He is, however, inconsistent in this. Often he quotes words which could have come to the area only after 1945 and, naturally, have more or less preserved their Standard Polish form (*Tomsk*, *front*, *ukraiinka*, etc.). He mentions that the area he studies was a crossing point for many influences (9). In several cases he refers to possible Czech influence, more rarely to Sorbian or Great Polish, and still more seldom to German. As a rule he does not indulge in confrontations of this kind at all. Interesting parallels may be drawn with Slovak or its dialects (in preserving the preposition *s* [p. 89], in alternating *x::ś* in the nominal declension [p. 98]). The change *o*>*y*, unclear to Bąk (47), could be elucidated from the better-known development of this type in some Northern Ukrainian dialects. The appearance of *i* before palatalized consonants (104) is paralleled in the Ukrainian Bojko dialects. But there are no Slovak and Ukrainian parallels in Bąk's study.

Thus, Bąk's book is basically devoid of generalizations and systematization. Moreover, it lacks the elementary arrangement into phonemes and allophones. True, these terms occur once in the book (96), but nowhere else does the reader know whether he is dealing with phonemic distinctions or nuances of individual pronunciation. And even these nuances are sometimes described more than vaguely ("Sometimes something like *å* is pronounced"! p. 48).

Bąk's historical commentary is not always reliable. When he derives *pórćyᵉńik* 'weaver' from *płótno* 'linen' (93, 115), it implies that Common Slavic **poll-* did not undergo metathesis in this particular word, which is fantastic. In fact, the underlying root is **pŭrl-* (cf. Russian *porty* 'clothes', Polish *parł* 'linen, sackcloth'). **lŭrl-* in these dialects seems to give *łort* in general, cf. *łornovec*, *na łorku*, etc., as quoted by Bąk, 36. There is nothing surprising in the local from *slf'oro*, instead of the form which Bąk expects: *cforo* 'four' (99). The latter would be, of course, an exact correspondence to Standard Polish *czworo*, but the form *slf'oro* is an accurate continuation of Western Common Slavic **čĭtvero*, according to the sound laws of the dialect (*c*>*s* before a dental stop).

From the viewpoint of merely presenting the date concerning a disappearing dialect, Bąk's study may also be subjected to criticism. Since this is probably the last recording of the dialect, it would have been advisable to include the material collected by Bąk's predecessors, to make complete all the data available, as P. Rost did with Polabian in his *Sprachreste der Drawäno-Polaben*, particularly because some of the older records are not easily accessible to most students. Still more regrettable is that no characterization of Bąk's informants is given. (Will the third part, the texts, contain it?) In recording a vanishing dialect it is particularly important to know from whom the given facts originate.

Nevertheless, one must be grateful to Bąk even for the mere collection and presentation of data, since otherwise they would certainly be lost for linguistics without leaving a trace. For all their unsystematic character, lack of adequate method in presentation, and lack of convincingness in explanations, these data are invaluable because of the very characteristics of the area. They are indispensable for the history of Polish inasmuch as Polish-Czech linguistic interrelations are concerned. Many a hypothesis would stand or fall in light of the facts presented by Bąk. I shall limit myself to three examples. (1) Against the predominant point of view, I advanced an assumption (*Word* 8.340f.) that depalatalization of consonants in such Polish words as *wesele*, *serce* cannot be deduced from Czech influence. Bąk shows that in Low Silesian dialects the consonants in those words still are palatalized (88, 103). It confirms my theory: had the depalatalization resulted from Czech influence, it would affect Silesia first of all. (2) In another article (now is press in *Homenaje*

Martinet, I) I suggested that the passing of *s* to *z* in Polish in such words as *dzwono* was in connection with *s*>*z* as a preposition. In the dialects described by Bąk there is no change *s*>*z*, and, significantly enough, they did not develop an affricate in *dzwono*, etc. (3) Polish scholars usually hold that Polish penultimate stress was preceded by a period of initial stress, which is still preserved in Czech: most Czech scholars, conversely, contend that Czech initial stress was preceded in Czech by penultimate stress, which is still maintained in Polish. Bąk quotes several examples in which the penultimate syllables have lost their vowels: *rozmaite*<*rozmaite*, *żoųxa*<*ževuxa*, *xaųpa*<*xalupa* (73, 85, 116). This seems to corroborate the view of Polish scholars that the initial stress is more archaic than the penultimate.

Columbia University. GEORGE Y. SHEVELOV.

MIKOŁAJ RUDNICKI, *Językoznawstwo polskie w dobie oświecenia.* (Poznańskie Towarzystwo Przyjaciół Nauk, Prace Komisji Filologicznej, XVII, 2). 33 pp. Poznań, 1956.

Rudnicki sets as the epoch of the Enlightenment in Polish linguistics the years 1741-1839. Yet time and again he encroaches upon the period later and cites texts of 1845, 1857, 1912... Within the epoch of the Enlightenment, as he defines it, he quotes indiscriminately the representatives of this trend and their opponents, the bearers of Romantic ideas, e.g. Mroziński. The pamphlet as a whole is supposed, according to Rudnicki himself, to be devoted to the subject of general linguistics in Poland of the period. Instead of this, it is made up of chaotic paragraphs treating problems of Polish grammar, Polish grammatical terminology, studies in Sanskrit, etc., with an admixture of general considerations on what is language and what are the national languages in particular, with a great many patriotic digressions.

The basic conflict between the Enlightenment, with its conception of universal nature of language based on logic, and Romanticism, for which logic is incompatible with language and which seeks and sees behind any language the "national spirit" in all its irrationality, is reflected in numerous quotations cited by Rudnicki. From these quotations one may deduce that in Poland there scarcely were pure representatives of the Enlightenment in linguistics. Most of them, despite their philosophy, promoted national slogans ("The very name of Poland and Poles shall exist

as long as the Polish language does", says Kopczyński), and did not stop short of writing "On the Spirit of the Polish Language" (the title of a pamphlet by Kopczyński, 1804). The reason for this probably was, besides the political subjugation of Poland, the lag in the penetration of Western European cultural trends to the Slavic countries. The trends which fought and mutually excluded each other in Western Europe sometimes reached Poland simultaneously and, therefore, coexisted in the writings of the same author.

Thus, certain conclusions may be drawn from Rudnicki's material. His own generalizations do not aim at characterizing the epoch under consideration but at finding in the casual utterances of Polish linguists of that time thoughts which are current in what Rudnicki calls "modern linguistics," for him represented by the names of Wundt and Rozwadowski! All in all, his pamphlet looks like a collection of slips with unsystematic excerpts made when perusing old books in a library. It cannot become a chapter in some future history of linguistics in Poland.

Columbia University. GEORGE Y. SHEVELOV.

ZDZISŁAW STIEBER, *Rozwój fonologiczny języka polskiego* (Structural Development of Polish). 96 pp. Warsaw, 1952.

With certain reservations, Stieber's book on diachronic phonemics of Polish can be taken as a quite typical example of today's linguistics in Europe where the emphasis remains on "process" rather than on all the problems and finesses of synchronic "arrangement".

Both the title (Structural Development...) and the arrangement of Stieber's book are, as was said, quite typical of modern European attitudes. In many ways it follows the traditional historical approach in presenting the development of Polish sounds in chronological order, but the shift in emphasis from "sounds" to "structure" opens many new views. The historical changes are analyzed not merely as articulatory or acoustic phenomena but in context with the whole. Their functioning in the system is the main point.

In Polish linguistic literature this is a novelty. A descriptive analysis of Polish Phonemes was given twenty years ago by S. Szober in his *Polish Grammar* (*Gramatyka języka polskiego*,

Warsaw, 1936, vol. II, pp. 54-89), but in historical studies structural considerations were rarely take up.

Encompassing a period of approximately one thousand years, Stieber first outlines the main features of the Polish vocalic system as inherited from the Common Slavic, and then traces its evolution up to our days. Sonants and consonants are treated in separate chapters. An additional section on Kashubian may contribute to a better understanding of today's somewhat complicated relationship between Kashubian and other Polish dialects.

In the chapter on vowels the author's "symmetric" inclination can be observed, especially in his treatment of the semi-vowels *ꙋ (u)* and *ь (i)*. As it is known, this has been a favorite consideration with many modern linguists who in a tendency towards symmetry in the pattern of correlations see one of the main factors in linguistic change. In Stieber's picture the two semi-vowels *ꙋ* and *ь* find themselves outside the regular pattern of correlations (long: short) at the time when this became the most striking feature of the system. Being "odd" they either had to become identified with other full vowels or reduced to zero. Stieber does not press the point of symmetry too strongly, as he probably could have done, in the case of nasals. He simply presents, quite convincingly, the shift from the notion of quantity to quality, at the time when the original qualitative difference between the two nasals *o* and *ę* (which in the period of quantitative dominance had got lost) was re-established. As for the problem of whether the vowel *ę* after the 16th century ceased to be a phoneme, it can only be said here that the reviewer dos not quite agree with the author, neither in principle (the theory of "facultative phonemes") nor in detail. (I discuss this problem in my article, "The Problem of Polish Phonemes", *Scando-Slavica*, Vol. II.).

The chapter on consonants outlines very neatly all the main phases of the evolution. A certain shift in the emphasis could have been elaborated, especially the role of the palatal: non-palatal correlation. Here again Stieber introduces the formula of "facultative phonemes" (k', g'), although with some reservations ("perhaps one should...")[1]. On the whole, Stieber rightly observes that the system of consonants did not change as much as the

[1] In his interesting remarks on the book («Uwagi na marginesie Rozwoju fonologicznego języka polskiego») published additionally in the *Proceedings* of the Societas Scientiarum Lodziensis, Linguistic Section, Vol. I, pp. 115-119, Stieber speaks of *k', g' as* «facultative phonemes» without the previous reservations (p. 117).

system of vowels. Theoretically it might have been of interest to speculate to what extent certain distributional disturbances leave the system open to further changes. But, of course, Stieber's book does not speculate on this point. It is primarily a concise account of the main historical facts. One of its additional merits is that it is skilfully documented with excellent source materials, some of which have never before been used for such purposes.

University of Wisconsin. ZBIGNIEW FOLEJEWSKI.

ZENON SOBIERAJSKI, *Gwary kujawskie* (The Dialects of Kujawy). Poznańskie Towarzystwo Nauk, Prace Komisji Filologicznej, Vol. XIV, fasc. 2, iv+128 pp., 1 map. Poznań, 1952.

Dialectology has always been a well represented branch of Polish linguistics. The two main schools—that of Kazimierz Nitsch, Cracow, for many decades the greatest authority on Polish dialects, and that of Witold Doroszewski, Warsaw, who modernized Nitsch's methods by putting greater emphasis upon statistics of frequency—complemented each other, and brought many interesting results. But there are still areas in Poland that require detailed field research, and, in spite of all post-war difficulties Polish linguists are busily filling the gaps.

The author of the book under review, Zenon Sobierajski, started his investigation of the Kujawy dialects some time ago, in connection with his dissertation on the language of Jan Kasprowicz, a poet of peasant origin, hailing from Kujawy, and in whose poetry the Kujawy dialect had a specific artistic function. This reviewer, who some ten years ago investigated the problem of the artistic function of dialectisms in literary works (cf. *La fonction des éléments dialectaux dans les œuvres littéraires*, Uppsala, 1949) is acquainted with Sobierajski's earlier work, *Elementy gwarowe w utworach Kasprowicza* (Poznań, 1950) of which the present study is an extension and, at the same time, a more precise linguistic treatment.

Sobierajski undertook extensive field research in 1946-1947 covering 87 villages representative of the Kujawy area (Middle-North-West part of Poland). His task was not easy, because the Kujawy area has always been somewhat transitional in character (lying between Great Polish and Masovian dialects) thus showing considerable heterogeneity in important features of both phonology and morphology. Consequently, Sobierajski's final picture, as

illustrated on his map, discloses quite a complicated net of isophones and isoglosses, concentrated in the South-Western part of the area. The main line of division is said to run East-West (to this reviewer, it looks as if it ran from the North-East to the South-West).

Especially intricate is the treatment of the nasal vowels (pp. 22-28) and also of the problem of the assimilation of consonants (pp. 44-49). In connection with this latter point, it should be remarked that we nowhere get a clear picture of the author's methods either in his recording of the data nor in his transcription of them. In some cases it is not easy to say whether the author neglected to analyze all the distinctions in separate sounds in detail or whether we may have individual differences in pronunciation. A similar doubt arises concerning the assimilation of *v* to a voiceless environment (pp. 31-32). The author notes only complete unvoicing or retention of voiced character, and does not mention the possibility of partial change.

University of Wisconsin. ZBIGNIEW FOLEJEWSKI.

LESZEK MOSZYŃSKI, *Geografia niektórych zapożyczeń niemieckich w staropolszczyźnie*. Poznańskie Towarzystwo Przyjaciół Nauk, Wydział Filologiczno-Filozoficzny, Prace Komisji Filologicznej, Vol. XV, fasc. 3, 94 pp. Poznań, 1954.

Moszyński studies the geographical distribution of German loan-words in Old Polish. Out of several hundred loan-words the author selects 16 as having distinct geographical limitations. His statistics, somewhat meager to be sure, corraborates of course what has been known regarding the main directions of German penetration into Polish. Masovia's loan-words are of *niederdeutsch* or *mitteldeutsch* origin. The second group of German loan-words (Great Poland and Little Poland) is of *hochdeutsch* or *ostmitteldeutsch* origin. While the material in this group shows some similarity to Old Czech, the Masovian group does not have any such formations.

Z. FOLEJEWSKI.

REINHOLD OLESCH, *Zur Mundart von Chwalim in der früheren Grenzmark Posen-Westpreussen*, Akademie der Wissenschaften und der Literatur, Abh. der geistes- und sozialwiss. Klasse, 1956, Nr. 6. 28 pp. Wiesbaden: Franz Steiner, 1956.

Chwalim, a village in the former Posen Province of Imperial Germany, was left on the German side of the 1919 border. The local Slavic vernacular was rapidly yielding before German, and, after the displacements that have taken place since 1945, it may be considered extinct. The inhabitants used to designate themselves as "Wends" and were accordingly believed to speak a Sorbian dialect. This was disproved by competent observation: the language of Chwalim was Polish, but a lower Silesian variety of it brought into contact with Great Polish by 17th Century Protestant settlers. Olesch makes available Wenker's questionnaire as filled in, in the late eighties, by the local schoolteacher with the help of his charges; he appraises its value as linguistic evidence and points out the features that are relevant for the identification of the dialect.

ANDRÉ MARTINET.

Kelica. Materialien aus dem Ketischen oder Jenisseiostjakischen. Aufgezeichnet von Kai Donner, bearbeitet und herausgegeben von Aulis J. Joki. (=Mémoires de la Société Finno-ougrienne 108). 134 pp. Helsinki 1955.

In addition to the work of K. Donner and A. J. Joki, as indicated in the title, this book also contains a Ket text originally collected by the ethnographer Hans Findeisen, as prepared for publication by Ernst Lewy. The book therefore combines the contributions of four men, all of whom have previously dealt with the culture or language of the Kets and two of whom (Donner and Findeisen) have carried on fieldwork among them.

The Kets (the older appellation, Yenisei-Ostyaks, has been abandoned in the Soviet Union) are the smallest and only western paleo-Siberian group. Their number is believed to be between 400 and 500 and is steadily decreasing because of Russian (and formerly Ostyak-Samoyed [Selkup]) acculturation. The only language known to be related to Ket is Kot which is now extinct. G. J. Ramstedt, Donner, Lewy, the sinologist W. Simon, Findeisen, Joki and others have expressed the questionable opinion (or conviction) that Ket and Kot are related to "Indo-Chinese". (No such avowal, however, is made in the book under review.) The first and only thorough presentation of the language was made by M. A. Castrén (died 1851), a sketch of Ket by E. N. Karger appeared in *Jazyki i piśmennosť narodov severa* III (1934), and the Kot verb has been discussed at length by K. Bouda (*Abhandl. f. d. Kunde des Morgenl.*, [*ZdMG*], 1935).

Kai Donner (1888-1935) was a specialist in Southern Samoyed languages who visited Ket settlements during his travels in Siberia. In addition to his fieldwork, he had occasion to work with a gifted, young Ket informant in Helsinki during the summer of 1928. The text and practically the entire glossary contained in *Ketica* are results of work with this informant who learned to use one of Donner's transcriptions.

Ketica embraces the last remains of unpublished information about the Ket language outside the USSR (except, perhaps, for a few fragments which may still turn up among Donner's papers). Aside from Findeisen's text, it contains a Ket-German glossary of somewhat 2000 entries (the first Ket dictionary since Castrén's) and another text with a German translation. Findeisen's text appears in two complementing variants and in his own transcription, as prepared for publication by E. Lewy. Donner's text is given in two transcriptions, his own and that of his informant. In the glossary, the main word in the entry is given in Donner's "broad (gröbere)" transcription, while the Ket words within the entry appear in Donner's highly detailed phonetic transcription. (The latter contains at least 44 vowel signs.) We are thus dealing with four transcriptions: Donner's broad one, Donner's narrow one, Donner's informant's, and Findeisen-Lewy's. Aside from the last of these, all are based on the speech of one informant (with a few exceptions indicated by the editor). This variety was unavoidable from the viewpoint of the faithful and conscientious editor, but its drawbacks will be obvious. This should be added to the fact that Donner's transcription in *Ketica* differs from both his own (other) broad transcription in his *Ethnological Notes About the Yenisey-Ostyak* (*Mém. Soc. F.-ougr.* 46) and from the transcription used in the USSR (Karger, *op. cit.*). None of the four transcriptions in the book are phonemic, though it would perhaps not have been impossible to derive an at least provisional phonemic transcription from Donner's detailed phonetic one. The most interesting point in this connection is the native Ket's use of Donner's broad transcription.

The glossary contains well-chosen paradigmatic, and other examples and thorough cross-references enable the reader to gain certain notions about Ket derivation. The cultural frame of the language is taken into consideration, so that the glossary may also be of value to the ethnographer. Russian loans and, to some extent, Tungusic and Ostyak-Samoyed ones, are indicated. The texts are difficult to analyze and were never quite clear to the

collectors, which made the editors' tasks insurmontable. This reviewer would have preferred an interlinar translation, no matter how faulty. Especially noteworthy are the pains the editor took to decipher the seemingly difficult manuscript, to add cross-references and references, both in the glossary and in the texts, and to provide the book with a German-Ket index and a useful, interesting introduction. The Société Finno-ougrienne which usually supports the publication of Uralic and Altaic works deserves special credit for sponsoring this important non-Uralic, non-Altaic source.

Abasiri (Japan). ROBERT AUSTERLITZ.

LAURI KETTUNEN, *Über die Wortfolge im Finnischen und Urfinnischen.* (=Annales Academiae Scientiarum Fennicae, Ser. B, Tom. 81: 3.). 67 pp. Helsinki 1953.

The specific question under discussion is: if the predicate of a Finnish sentence is modified by an adverb or by an adverbial phrase and if this adverb or adverbial phrase stands at the beginning of the sentence, the order of subject and predicate may vary. E.g., *Eilen isä tuli kotiin* 'Father came home yesterday' (*eilen* 'yesterday', *isä* 'father', *tuli* 'came') as against *Eilen tuli isä kotiin.* Some Finnish authorities have maintained that inverted order, i.e., predicate before subject, (1) is wrong, (2) is bad Finnish, (3) serves to express emphasis, (4) serves to express what they call "species" (indefiniteness as against definiteness), etc. The author, formerly professor of Estonian and cognate languages in Helsinki, first polemizes against all of these views and then adduces examples from Finnish, three Carelian dialects, Vot, Estonian, and Livonian to show that inverted order exists in all of these languages and existed in Proto-Baltic-Finnic. His arguments are astutely presented and well documented.

Abasiri, Japan. ROBERT AUSTERLITZ.

SØREN EGEROD, *The Lunglu Dialect, A Descriptive and Historical Study of a South Chinese Idiom.* xvii+284 pp. Ejnar Munksgaard, Copenhague, 1956.

L'ouvrage est une thèse soutenue à l'Université de Copenhague. Il s'agit d'une monographie du parler d'une chinoise de Lungtu

(Long Tou) à 50 kilomètres au Nord de Macao. Des textes divers, récits, contes, chansons, dans le langage courant, et deux textes en prononciation littéraire sont publiés (118-207). Un vocabulaire extrait des textes et complété par la prononciation des caractères utilisés par Karlgren dans ses travaux linguistiques, est classé par rimes (85-117). L'étude phonologique du dialecte (27-51) et une esquisse grammaticale (52-70) complètent l'étude synchronique.

L'intérêt de ce dialecte réside dans ses origines : parlé par les Hoklo, colons originaires du Fou-Kien, il conserve les traits des dialectes Min de cette province, qui ont échappés à l'unification de la période des T'ang ; il présente comme ceux-ci une prononciation différente des caractères dans la conversation et dans la lecture. Installé en territoire cantonnais, ce dialecte a été influencé par celui de la ville voisine de Shekki (71-82).

L'auteur élabore une notation cohérente pour tous les dialectes dont il sera question dans la partie historique ; Cantonnais de Hong-Kong, de Shekki, Pékinois, et dialectes Min de Swatow, de Amoy et de Fou-Tchéou (7-26). La partie historique consiste en un classement des mots recueillis groupés selon les restitutions de Karlgren : l'examen des initiales occupe les pages 212 à 226, celui des finales les pages 228 à 271 et enfin les tons sont étudiés dans les pages 272 à 279. Un résumé de deux pages, en anglais et en danois, termine l'ouvrage.

On ne peut qu'être reconnaissant à l'auteur pour son important travail et pour des matériaux qu'il nous fournit. On regrettera, certes, que la partie historique n'ait pas été élaborée d'une façon plus structurale. Mais il faut reconnaître que cela l'aurait entraîné trop loin. Malgré les rapprochements avec les autres dialectes connus, la densité des points étudiés n'est pas suffisante : il nous faudrait un véritable atlas linguistique du Sud-Est de la Chine, pour élucider l'histoire de ces dialectes dont l'évolution « normale » a été constamment troublée, soit par l'influence de la langue officielle de communication administrative, soit par les mélanges dialectaux résultant des mélanges de populations. On sait que, dans ces régions, les populations trop denses, chassées par les famines ou les troubles sociaux, émigrent vers le Sud et forment une véritable mosaïque.

André G. Haudricourt.

A. Capell, *A New Approach to Australian Linguistics*. Oceania Linguistic Monographs, No. 1. 103 pp. University of Sidney, Australia, 1956.

LOUIS J. LUZBETAK, *Middle Wahgi Phonology*. Oceania Linguistic Monographs, No. 2. 48 pp. University of Sidney, Australia, 1956.

Les deux monographies qui nous arrivent d'Australie sont aussi différentes que possible par leurs méthodes et par leurs buts.

Depuis presque vingt ans, A. Capell dirige et inspire les recherches linguistiques en Australie. Il nous donne ici une espèce de compendium de l'Australien, un peu à la manière dont Bopp et Schleicher nous parlaient de l'indo-européen, il y a un siècle. Et ceci est très intéressant : toutes les langues de ce continent sont apparentées (sauf celles de la péninsule d'York, proches de la Nouvelle-Guinée) ; il est possible de reconstituer le vocabulaire commun, de comprendre l'évolution de la structure grammaticale qui aboutit dans certaines régions à des langues à préfixes et dans d'autres à des langues à suffixes. Mais, en fait, le lecteur reste sur sa faim, car il y a très peu de descriptions satisfaisantes des langues particulières d'Australie. Les publications du fils Strehlow ont montré, par exemple, que l'Aranda avait été très mal décrit, de sorte que tout le travail d'analyse qu'avait fait Sommerfelt est à reprendre. On souhaiterait être mieux renseigné ici. Il est à craindre que lorsque Capell ne restitue, pour l'Australien commun, que trois voyelles, *i*, *a*, *u*, il agisse de la même façon que Bopp pour l'indo-européen.

Au contraire l'ouvrage de Luzbetak établit le système phonologique des langues Papoues des Montagnes de Nouvelle-Guinée d'une manière très satisfaisante et selon les meilleures méthodes de la linguistique contemporaine ; il s'agit de langues à registres, d'un consonantisme assez pauvre, comportant des demi-nasales, comme les langues mélanésiennes.

ANDRÉ G. HAUDRICOURT.

IMPRIMÉ PAR A. BONTEMPS, LIMOGES (FRANCE)

Dépôt légal : 3e trimestre 1957